UNIVERSITY OF MICHIGAN
JUL 13 1955
SOCIAL SCIENCE LIBRARY
BF
1
.J86

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

A Method of Controlling Stimulation for the Study of Space Perception: The Optical Tunnel: J. J. GIBSON, J. PURDY, AND L. LAWRENCE 1

The Role of Simultaneous Contrast in Brightness Constancy: H. LEIBOWITZ, N. A. MYERS, AND P. CHINETTI 15

Effect of Photometric Brightness on Judgments of Distance: J. COULES 19

Development of Response Generalization Gradients: C. P. DUNCAN 26

Avoidance Learning to the Onset and Cessation of Conditioned Stimulus Energy: G. B. KISH 31

Studies of Distributed Practice: XIII. Interlist Interference and the Retention of Serial Nonsense Lists: B. J. UNDERWOOD AND J. RICHARDSON 39

The Role of Irrelevant Stimuli in Human Discrimination Learning: M. HAMMER 47

Two Tests of the Sheffield Hypothesis Concerning Resistance to Extinction, Partial Reinforcement, and Distribution of Practice: W. WILSON, E. J. WEISS, AND A. AMSEL ... 51

Resistance to Extinction as a Function of the Type of Response Elicited by Frustration: H. M. ADELMAN AND J. L. MAATSCH 61

Effects of Sequential Dependencies on Instrument-Reading Performance: V. L. SENDERS AND J. COHEN 66

Variations in Sending Morse Code: C. WINDLE 75

American Psychological Association

Vol. 50 No. 1 July 1955

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1955 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright, 1955, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 50, NO. 1 JULY, 1955

A METHOD OF CONTROLLING STIMULATION FOR THE STUDY OF SPACE PERCEPTION: THE OPTICAL TUNNEL[1]

JAMES J. GIBSON, JEAN PURDY, AND LOIS LAWRENCE

Cornell University

The apparatus to be described is designed for the study of the stimulus conditions which yield tridimensional visual perceptions. The methods of psychophysics can probably be extended to the investigation of space as well as color if devices can be built for the control and systematic variation of these stimulus conditions. Devices have been constructed in the past for the control and systematic variation of luminous intensity and wavelength; we now need devices for the control and systematic variation of luminous *pattern*.

The stimulus for visual perception is focusable light, that is, light which is capable of forming a pair of images in the two eyes. In an environment of clear air (which transmits light) and solid surfaces (which reflect light) these images will constitute a sort of dual projection of the solid portion of the environment.

The retinas are adapted to register this dual projection. The receptor system is sensitive not only to the pattern of each image but also to the difference between the patterns (their disparity) and to the change of both patterns in time (their motion). Each pattern is composed of *transitions*, or differences in luminous energy, and we know that the eye functions by accommodation so as to make these transitions as abrupt as possible, that is, to maximize what is called the "sharpness" or "definition" of the image. In the focused image, the abrupt transitions will generally correspond to the edges or margins of objective surfaces. Gradual or slow transitions, if present, will generally correspond to such physical conditions as the penumbras of cast shadows or to the shading on curved surfaces in the environment.

Within the gross areas of relative light and dark in the image there will usually be found finer areas. There may exist *alternating* transitions of

[1] This research was supported in part by the United States Air Force under Contract No. AF33(038)-22373 monitored by the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center, Lackland Air Force Base. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part, by or for the United States Government. In addition to the contributors listed, the work was aided in its earlier stages by Thomas Marill.

luminous energy along both meridians of the image, or "optical texture."[2] The optical texture of the image is generally a projection of the physical texture of the reflecting surface, that is, the pattern of one corresponds geometrically to the pattern of the other. When the density of the transitions along one meridian exceeds the density along the other, this "one-way compression" of the texture (**2,** p. 173) generally corresponds to the *slant* of the surface to the line of sight. A continuous increase in the density of such transitions along one dimension of a two-dimensional image corresponds to the *recession* of the surface, for example, a floor or ceiling. The density of the texture surrounding a given point in this image corresponds to the *distance* of the corresponding point in the surface. Such facts of optical geometry help to resolve the paradox of how a retinal image in two dimensions can yield information about a space of three dimensions.

The *density variables* of the monocular image are associated with physical distance by virtue of the geometry of perspective, which is concerned with a projection to one center at one instant. The visual stimulus, however, is ordinarily a projection to two centers, and moreover each center of projection ordinarily changes with time. The image in one eye differs from that in the other, and each image is altered whenever the observer changes his position. Hence we must also look for variables in the visual stimulus which depend on the geometry of *parallax* or, more generally, the geometry of *projective transformations*. If the visual stimulus is a dual, moving projection, the density variables in the array will be supplemented by variables of disparity and deformation. Along with degrees of density along a meridian, the array will manifest degrees of displacement of one image relative to the other, and degrees of displacement in each image relative to what it was before. A reasonable hypothesis is that the latter can be analyzed as gradients of the array and considered stimulus variables for visual experience.[3] The angle of inclination of a textured surface to the line of sight, for instance, is ordinarily specified by three concomitant gradients in the stimulus array: the rate of change of density, disparity, and motion of the texture elements.

A Method of Presenting Spatial Stimulation

In order to discover whether these gradients actually serve as stimuli for impressions of a spatial world, a means must be found of experimentally producing and controlling them. For this purpose a device has been constructed which delivers to each eye a sheaf of light rays whose cross section is a set of concentric rings of alternating high and low intensity ("white" and "black"). It consists physically of a series of

[2] The senior writer has suggested elsewhere (**2,** Chap. 5) that the elements of an optical texture might be defined as "spots" with "gaps" between them, or as "cycles" of intensity each composed of a spot and a gap. This definition, apart from its mathematical vagueness, now seems to be an error. The element of optical texture is best defined as an abrupt change in luminous intensity, either an increase or a decrease. These changes themselves, not the areas they segregate, seem to be what is of first importance in the retinal image. They can be analyzed mathematically for slope and rate of change of slope. The retinal image, in short, is not composed of objects but of transitions. The "rays," in terms of which we analyze the optical stimulus, should be conceived as points of change, not as entities.

[3] Introspectively, these two modes of optical stimulation are manifested as the illusory double images of objects in the binocular field and the illusory apparent motion of objects in the field. These impressions can be called sensations and the apparent displacements can be called clues or cues for assumptions or inferences about the distances of the objects, but this line of theorizing is obviously roundabout.

circular holes in alternately black and white diffusing surfaces hanging behind one another, perpendicular to the line of sight. The surfaces are sheets of plastic uniformly illuminated from below. A longitudinal section of the resulting sheaf of light rays is diagramed in Fig. 1. The solid lines represent reflecting surfaces and the broken lines represent divisions between lighter and darker ray sheaves, that is, abrupt increases and decreases in luminous intensity. The end of this pseudo tunnel consists of a white or black sheet, not shown, which fills the center of the stimulus array.

Only the stimulus for one stationary eye is represented in Fig. 1, and the eye is centered with the tunnel. The cross section of its ray sheaf is shown in Fig. 2A. The stimulus for another eye (or for another position of the same eye) must be represented by an off-center projection. The cross section of such a ray sheaf is shown in Fig. 2B from a point to the *right* of center. In the latter picture it may be noted that the peripheral-to-central rate of increase of density is shallower on the left and correspondingly steeper on the right, but is unaltered downward or upward. It may also be noted that the over-all density of this optical texture is very coarse.

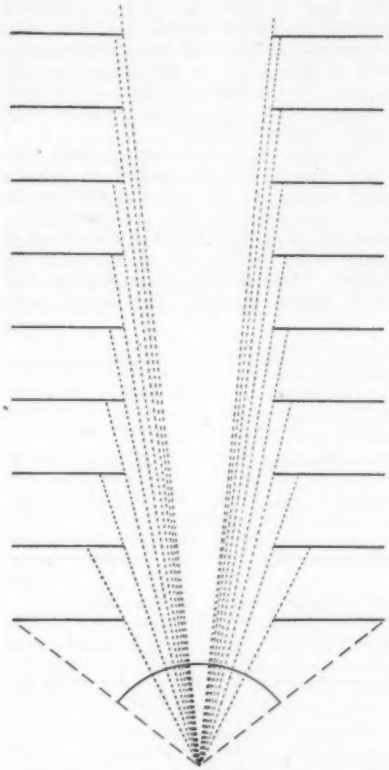

FIG. 1. Longitudinal section of an optical pseudo tunnel. Nine elements or transitions are shown as projected to a single centered eye. The increase in density of transitions from periphery to center of the array is evident on the angular cross section.

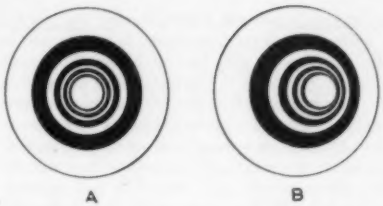

FIG. 2. Perspective cross sections of the optical tunnel of Fig. 1. Transitions are shown as white to black or the reverse. The picture on the right (B) represents a projection to a point to the right of the centered eye.

The series of apertures shown in Fig. 1 is one of a variety of arrangements which can be made. The size of the first aperture and its distance from the eye determine the *angular extent* of the stimulus array, that is, the amount of the field of view filled by it. The length of the pseudo tunnel determines to what extent the rings fill up the array toward its center, that is, the *maximum* density of the texture, a density which can increase up to a theoretical vanishing point as the tunnel lengthens. The number of sheets placed in a tunnel determines the *mean* density of the array, that is, the total number of texture elements in it. In other words, the physical spacing of the apertures determines the optical variables of the stimulus. The artificial visual environment produced is "flexible," in that it can be altered by E in the systematic way necessary for psychophysical experiments.

A peripheral-to-central increase of density is represented by the arrangement of Fig. 1 and 2. A constant density of texture or zero gradient can, however, be arranged. This corresponds to a surface of zero optical slant, i.e., a flat frontal surface. The arrangement is shown in Fig. 3. It should be noted, however, that the ray sheaf represented is single and stationary, and that if the rays are projected to *two* points or to a *moving* point, a peripheral-to-central gradient of disparity or of relative motion would appear in the array which would be inconsistent with the zero gradient of density. The two parallax gradients would thus be discrepant with

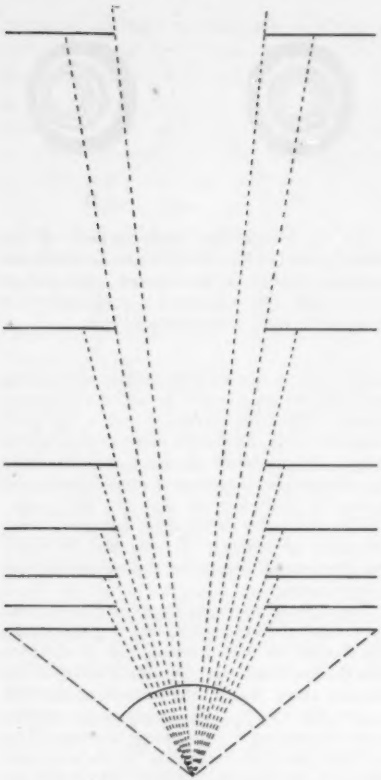

FIG. 3. Arrangement of a pseudo tunnel which provides a constant density of transitions from periphery to center.

the perspective gradient, and the result for perception could be observed.

If the apertures of the tunnel are of constant size, variations of the stimulus are produced by the spacing of the sheets. It is possible, however, to cut a series of holes of decreasing (or increasing) size. Variations of the stimulus then depend on *both* the size of the holes and the spacing of the sheets. A given projection to a stationary point can, geometrically speaking, arise from an infinite set of tridimensional arrangements and, accordingly, there is an infinite set of possible combinations of serial size and serial spacing which will yield a given pattern of concentric rings. Figure 4 illustrates a tunnel which is one such projective equivalent to the "normal" tunnel shown in Fig. 1.

It is important to realize that, although the gradients of texture density in these projective equivalents are identical, the gradients of disparity and of relative motion are not. The latter are not as steep in Fig. 4 as they are in Fig. 1. Consequently it might be predicted that the percepts induced by these two tunnels would be the same with monocular vision and a motionless head, but different with either binocular vision or a mobile head or both.

The specifications for an optical tunnel.—The aim is to construct a synthetic perception of tridimensional space by arranging the differences of intensity in a stimulus array. Great care must therefore be taken to eliminate all variables in the light entering the eyes except the intended alternations of dark and light rings. To this end, the reflecting sheets must be thin and their surface must be matte and very smooth. Vinylite plastic has been employed. When a clean black or white sheet of this material is observed monocularly through a viewing tube, accommodation fails and the impression of a dark or light "film" at the end of the tube results. The surfaces, therefore, are diffusing reflectors which yield no visible texture when seen at a sufficient distance.

The holes cut in these surfaces must have sharp edges in order to produce an optically sharp margin and the cut must therefore be beveled on the invisible side. The plastic sheets used were .03 in. thick, which is thin enough to permit exact cutting but thick enough for rigidity. Material which warps, such as cardboard, cannot be used. The use of metal

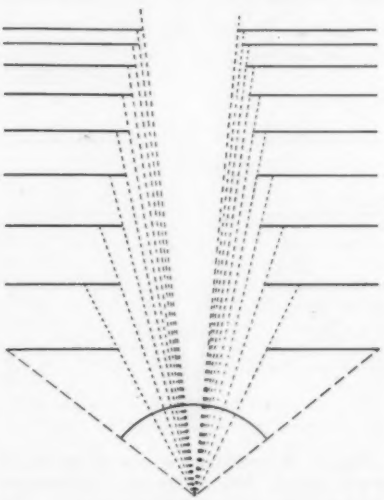

FIG. 4. Arrangement of an optical tunnel which is equivalent of that of Fig. 1 for the fixed single-point projection shown.

instead of plastic, it now appears, would probably be desirable. Circular holes are easier to cut than square holes.

The surfaces and holes in the existing tunnel are of rather large dimensions in order to minimize the effect of imperfectly smooth surfaces or imperfectly sharp edges. Exploratory work with a tunnel of small dimensions revealed this effect. The holes are 12.6 in. in diameter and the sheets are 36 × 42-in. rectangles. The maximum length of the existing tunnel is 21 ft. along one wall of a room. There is enough space for a 15-ft. comparison tunnel at right angles to the standard tunnel along the adjacent wall. The plastic sheets had to be large relative to the holes in order to make possible a long tunnel composed of only a few texture units, that is, relatively coarse concentric rings. Relatively fine concentric rings in a tunnel of equal length can always be obtained by hanging a large number of sheets behind one another. Fifty or more sheets can be set up within 21 ft.

The top of each sheet is sandwiched between two strips of aluminum angle whose ends rest on a pair of steel tracks 21 ft. long. These tracks, running the length of the tunnel, are supported at a height of 5.5 ft. from the floor, level, and parallel. The centers of the holes are thus at eye level for a seated *O*. The sheets can be spaced by sliding them along the tracks, or by adding to or subtracting from their number. The end of the tunnel can be set at any point short of 21 ft. by hanging a plain sheet at that point.

Uniform illumination of the alternating black and white surfaces which reflect alternating intensities of light had to be arranged if the different intensities were to depend wholly on the different reflectance values. The method adopted was to set a series of fluorescent lamps end to end underneath the plastic sheets, each tube overlapping the next slightly, so as to provide each surface with an equal amount of illumination. Reduced illumination was obtained by covering the lamps with one or more layers of paper.

The viewing position of *O* may be either free, or fixed by a headrest or biting-board. The headrest can be equipped with occluders which will permit either binocular or monocular vision. The whole device is screened from the view of *O* at all times so as to prevent any expectation of the kind of scene to be viewed. When ready to observe, *O* faces a rectangular sliding panel which can be raised and then lowered. The plane of this window is at 40 cm. from the eyes. At any convenient distance beyond it, the first of the series of apertures can be set.

The two-mirror setup for constructing an optical tunnel.—If a pair of mirrors of identical shape are set up facing one another in parallel planes, and a small hole is made in the silvering of one mirror to which an eye can be applied, a "normal" optical tunnel (Fig. 1) can be constructed of small bulk and at little cost. Such a tunnel is of indefinitely great length, i.e., it recedes to a vanishing point. Its elements have as much depth as the distance between the two mirrors and its over-all density can be varied by altering this distance. The edges of the two mirrors repeat themselves as successive reflections down what appears to be an endless corridor. The optical geometry is such as to produce the sheaf of rays in Fig. 1. The perception of distance is compelling. A sort of dark spot appears in the space of this corridor, which is the image of the viewing hole. A pair of holes or a slit can be made for binocular viewing, but these will be correspondingly imaged.

The size and spacing of the tunnel elements cannot be independently varied in this arrangement. It is doubtful if the elements can be made to consist of simple transitions of luminous intensity as can those of the aperture tunnel, but this is not certain. The walls will consist not only of the mirror edges but also of anything lateral to these edges. It is possible to fit a sort of translucent diffusing collar around the edges of the mirrors (circular or square) and to illuminate the tunnel through it. Textures can be imposed on this surface. Preliminary work suggests the necessity of using front-surface mirrors.

No formal experiments have yet been performed with this device. It is described here so that other *E*s may work with it if they choose.

Results

Does a Pseudo Tunnel Yield the Perception of a Phenomenally Real Tunnel?

The first question to be asked about the aperture tunnel is whether it can produce the experience of a solid objective environment. Can the light be so manipulated as to yield the perception of a continuous substantial surface receding from *O* in the form of a cylinder or tube?

Preliminary experiments suggested that when the stimulus array consisted of only a few margins from periphery to center, they looked like edges and the rings between the margins looked frontal to the line of sight; the tunnel was seen for approxi-

mately what it was. When, however, the array consisted of *many* margins, they no longer looked like holes and the rings appeared to be black-and-white stripes on the interior of a cylinder. It is possible, then, that a surface is seen when the jumps of luminous intensity in the array (Fig. 2) are sufficiently dense.

In order to verify this hypothesis, three arrangements of the apertures were set up in succession so as to give three degrees of density in the stimulus: a 7-element tunnel, a 13-element tunnel, and a 19-element tunnel. In the first there were 7 sheets 90 cm. apart, in the second 13 sheets 45 cm. apart, and in the third 19 sheets 30 cm. apart. The total length of the tunnel was about 6 m. for all three arrangements, and the physical distance from the first aperture behind the square window to the last aperture in front of a solid sheet was always 540 cm. These three arrangements were shown to three groups of naive *O*s. They were simply asked what they saw. The sheets were adequately illuminated (about .1 ft.-candle), both eyes were used, and the headrest permitted some sideways movement of the head. The stimulation, in short, was not impoverished or reduced to a single fixed sheaf of rays; it was a dual fluid sheaf of rays.

The reports desired were not introspective descriptions but statements of what *O* perceived as "there." The instructions, accordingly, were as follows:

"When I lift the panel in front of you there will be something there that I would like you to describe to me. Tell me as much about it as you can, in as much detail as possible. I will first give you a very brief look, about 5 sec., and later, a much longer look. Note such things as color, shape, dimensions, etc. You may compare it to something you may have seen before if you like. I will ask you a few questions after you have reported. (Subsequent questions, using whatever noun *O* used in his report, were as follows.)

"(1) Would a ball, given a little push, roll from one end of the . . . to the other? (2) How far is it from you to the end of the . . .? (3) Is this . . . the same diameter all the way through, or does it narrow, or get wider at the other end? (4) What is the diameter of this end? Of the far end? (5) (If stripes or bands are mentioned) How are they arranged? How wide are they? Do they all look the same width, or do they get wider or narrower at the far end? (6) Where is the light coming from? (7) What does this look as though it were made of? How would you go about making it?"

Each *O*, in complete ignorance of the experimental setup, was led into a sort of booth in the experimental room, the window was opened for 5 sec., he peered into it, and his first report was recorded. All *O*s used nouns like "tube," "cylinder," "tunnel," "sewer pipe," or something equivalent, but whether the percept was that of a *solid continuous* surface had to be determined from the description and from the answers to questions (especially No. 1 and 7 above). On the basis of these records it was found that all percepts could be put into one of three categories: (*a*) a continuous solid cylinder, (*b*) solid cylindrical segments with space or air between the segments, and (*c*) apertures or holes in frontal surfaces. The second category consisted of a compromise between the first and third, that is, a percept in which one ring (usually the black) looked like a surface but the next (usually the white) looked like a gap. It may be noted that the third category is "correct" but this is irrelevant to the experiment.

The results are given in Table 1, in the form of percentages of the group of *O*s getting each type of percept. It can be noted that as the number of intensity jumps in the stimulus array increases, the impression of solidity or continuity of surface also increases. With as few as 19 elements the percept of a "real" surface is induced in 95% of the *O*s.

The results when *O*s were permitted to look into the pseudo tunnel for as long as they liked are given in Table 2. The instructions were probably such as to put *O*s on the lookout for

TABLE 1

SOLIDITY OF TUNNEL AS A FUNCTION OF DENSITY OF TEXTURE WITH SHORT EXPOSURE

O's Report	7-Card Tunnel (N =21)	13-Card Tunnel (N =20)	19-Card Tunnel (N =21)
Solid cylinder	33%	70%	95%
Solid segments with air between	19%	10%	0%
Series of apertures	48%	20%	5%
Total	100%	100%	100%

TABLE 2

SOLIDITY OF TUNNEL AS A FUNCTION OF DENSITY OF TEXTURE WITH PROLONGED EXPOSURE

O's Report	7-Card Tunnel (N = 21)	13-Card Tunnel (N = 20)	19-Card Tunnel (N = 21)
Solid cylinder	33%	45%	62%
Solid segments with air between	33%	10%	19%
Series of apertures	33%	45%	14%
Total	100%	100%	95%*

* The additional 5% is accounted for by one *O* who saw *part* of the tunnel as segments and the *remainder* as apertures.

an "illusion." The variety of objects seen by them became much greater, and the descriptions were equally various. The percepts could be classified in the same way as before, however, chiefly on the basis of Question 1. There was either a continuous solid surface on which a ball would roll, or there was no such surface, or there were places where a ball would drop out. With prolonged observation the frequency of reports of a solid cylinder decreased for the denser tunnels, but the trend of the results was the same as before.

The reason for the increased tendency to see edges and apertures with long exposure may be that stimuli for the perception of edges had not been wholly eliminated from the stimulus array and could be detected by some *O*s. Minute high lights reflected from the physical edges would be such stimuli. This possibility will be considered later.

It is notable that no *O* in this experiment ever saw as solid the space between the square window in front of his eyes and the circular hole in the first plastic sheet. This peripheral portion of the array was always perceived correctly as an air space. The edges of this window, of sawn wallboard, were perceptible as such. Between the first and second circular apertures, however, there was very frequently the clear perception of a black (or white) surface, as if a sort of stovepipe had been fitted to the hole in the surface. Only at this junction would an imaginary ball rolling toward *O* drop out. If this first ring was seen as solid, the entire tunnel appeared so. This perception was often wholly compelling, even for *O*s who knew the physical setup.

The solidity of the pseudo tunnel under the best conditions was also indicated by the reports of what it seemed to be "made of." Paper, cardboard, plastic, or metal were suggested, and the surface was said to be painted or to consist of sections joined together.

It is concluded that the pseudo tunnel will produce the optical stimulation for the perception of a solid surface in three dimensions. The critical factor for solidity seems to be density or frequency of the abrupt transitions between light and dark in the sheaf of light rays. As few as 19 jumps of intensity from periphery to center of the array will yield a surface perception. A subsequent arrangement where the elements were increased to 36 yielded the perception of a solid tunnel in all *O*s tested.

Perception of distance and size in the optical tunnel.—All subjects in the foregoing experiment perceived a three-dimensional scene. No one ever reported seeing a flat surface or a picture. The so-called cues of perspective, binocular disparity, and motion parallax were present and can be specified as gradients in the stimulus array. But the so-called cue of "known size" was absent. The scene was not identifiable as any well-known entity, even when it appeared most thing-like. It might be supposed that the unfamiliarity of this

TABLE 3

APPARENT LENGTH AND SIZE OF THE PHENOMENAL TUNNEL

(ACTUAL DIAMETER = 12.6 IN.; ACTUAL LENGTH = 17.7 FT.)

Dimension	7-Card Tunnel		13-Card Tunnel		19-Card Tunnel	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Length (ft.)	10.2	3.3	10.5	5.2	13.9	6.8
Near diameter (in.)	13.3	2.7	14.1	5.4	15.4	4.0
Far diameter (in.)	11.7	3.5	12.4	6.4	12.0	5.0

artificial space would make its size and its depth difficult to judge, and would make it as likely that the walls should appear to converge as that they should appear to be parallel. The *O* had no "information," apart from the light entering his eyes, as to whether the scene was a cone or a cylinder. Peripheral to the tunnel, there was visible, of course, a square window on a wallboard surface, part of a headrest, and *O*'s own nose, which is a familiar object in any normal field of view.

All *O*s were required to estimate in feet or inches the length of the tunnel, its near diameter, and its far diameter. They were also asked to say whether its sides were parallel. The mean estimates are given in Table 3 for the reports after prolonged inspection. The length was in general underestimated, but less so with the arrangement which gave the greatest solidity. Considering that these are absolute judgments by unpracticed *O*s, they are not far off. The diameters are also not far off, although apparent size seems to increase somewhat with solidity. Two-thirds of the *O*s reported that the sides were parallel, but one-third said they got closer together as they receded. The decrease in mean far diameter as compared with mean near diameter reflects this difference of opinion.

These results are exploratory, but it can safely be concluded that a phenomenal tunnel which approximates to solidity is seen in approximately its true scale. Both size and distance are perceived with some degree of accuracy and there is constancy of size with increasing distance.

Does the Phenomenal Tunnel Disappear When the Steps in Luminous Intensity are Eliminated?

The array of light reflected into the eyes from the optical tunnel consists theoretically of steps of intensity determined by the difference in reflectance between the two kinds of plastic, each being equally illuminated. A test can be made of the extent to which this theoretical condition is fulfilled. If surfaces of the *same* reflectance are substituted for alternating surfaces of different reflectance, the steps of intensity should disappear. The stimulus should then be homogeneous and the surface reflecting it should become invisible. The experiment predicts that an arrangement of physical surfaces in good illumination can be made to vanish.

The substitution of either all-black or all-white surfaces for the alternating ones provides a kind of control for the proper arrangement of an optical tunnel. If the microstructure of the surfaces is too coarse, if the illumination falling on each successive surface is not the same, or if the cut edges of the apertures are thick enough to yield even hairlines of reflected light or shadow, then the light entering the eyes will *not* be homogeneous. Brightness contrasts will appear in the field, differential

accommodation and convergence of the eyes will occur, and something like the edges and surfaces of an ordinary visual world will ensue. If, however, the array of light entering the eyes *is* homogeneous, or sufficiently so, none of these conditions will apply and the resulting impression will be a void, film, or fog (**4**).

The experiment was set up first with all-black surfaces and later with all-white surfaces, of 8 and 10 elements, respectively. The experiment showed the imperfections in previous setups. A considerable amount of trial and error was required to eliminate them. The presence of reflections or shadows from the cut edges of the sheets seemed to be the greatest cause of inhomogeneous stimulation. An arrangement was finally achieved with white sheets in which the holes had been cut with a "ring and circle shear" (a machine designed for sheet-metal work). For this setup the beveled edges could be made invisible on the upper semicircle of each aperture and almost invisible on the lower semicircle. Illumination was wholly from below, and there is reason to believe that a balanced illumination from above would have made even these disappear. A sufficient approximation to invisibility was obtained to warrant the conclusion that *when differences in luminous intensity within the stimulus array are made to approach zero, the objective or surface-like qualities of the percept tend to vanish.*

During the early stages of this experiment, the all-black or all-white tunnel appeared relatively insubstantial. It looked filmy, glassy, or transparent, and the interior seemed full of what was variously called smoke, haze, mist, or fog. This was a "dark" fog in one setup and a "light" fog in the other. With further adjustments and modifications the fog became thicker and the walls became less definite. When contrast between the surfaces had been nearly eliminated, it is fair to say that the tunnel as an object had practically disappeared beyond the first aperture. A few shadowy circles sometimes appeared within the first aperture, and a fleeting cylindrical impression then resulted, but the main impression was that of a film or fog.

In order to verify these observations, 10 *O*s were asked to describe what they saw when the observation window was raised, with both monocular and binocular vision. All had knowledge of the actual setup, and all had some training in visual observation. Words like filmy, translucent, soft, milky, hazy, or misty were generally applied to the luminous circular area. Sometimes this looked flat or two-dimensional, sometimes it looked deep, like "3-dimensional light," and for two *O*s it looked like a homogeneous convex sphere or disk which later became concave. Faint rings or circles were sometimes (but not always) apparent, and an impression of depth usually accompanied this. Something tunnel-like was often seen, but the reports were variable as between *O*s and from one moment to the next. There was little difference between the reports with monocular and binocular vision except for a somewhat greater tendency to depth in the latter case. The conclusion above seems justified.

The fact that a set of substantial objects can be made invisible by manipulating reflectance and illumination is not important in itself, however interesting it might be to a professional magician. An even more interesting question is why objects are usually *visible*. What the psychologist needs to know is the general relation between stimulation and per-

ception. The deeper implication of this experiment is that visual objectivity has a basis in stimulation. The previous experiment indicated that this basis may be found in the density of abrupt transitions in the stimulus. The fact that both experiments can be thought of as setting up an illusion or a misperception with respect to the physical sources of the stimulation is incidental, and can have misleading implications. The arrangements were not objects of perception, but were *E*'s devices for systematic variation of the stimulus.

The implication of these first two experiments, then, would be missed by calling them illusions or emphasizing that *O*s were deceived or misled. The construction of visual illusions for their own sake, however interesting, does not constitute systematic research. The optical pseudo tunnel is comparable to some of the striking illusions of space devised by Ames (**5**), but its purpose and implications are different. It enables *E* to produce synthetic perceptions of space so as to test hypotheses about the natural perception of space.

Fixed Monocular Vision as Ambiguous Stimulation

The experiments so far described permitted *O* binocular vision and head movement when looking into the pseudo tunnel. The stimulus consisted of a dual ray sheaf and was "fluid" instead of fixed. The question which arises is what happens to perception when the stimulus variables are restricted to those of a single ray sheaf to a fixed point such as is represented in Fig. 1.

A sheaf of rays to a fixed point may have for its reflecting source any surface which reflects that particular sheaf to that point. All such possible surfaces are related to one another by geometrical transformation. They are members of a mathematical *family* of objects. The fixed single ray sheaf, then, may be called *ambiguous* with respect to the different objects of the family. Applying this rule to a tunnel-shaped surface, it says that the fixed projection illustrated in Fig. 1 might be reflected from the interior of a regular striped cylinder, or from a longer tube whose walls diverge as they recede (like a megaphone), or from a shorter one whose walls converge as they recede (like a cone such as Fig. 4), or from a plane surface frontal to the eye (like the picture in Fig. 2A), or finally from the *exterior* of a short conical surface protruding toward the eye. The cylinder with parallel sides, the "tunnel," is only one of an infinite set of possible surfaces corresponding to the fixed monocular stimulus. Considered as an object in the world it is perhaps more "probable" than the others, but it is not specified in the projection. It is true that if the stripes of the regular cylinder are *equally* spaced, the stripes of all the other possible surfaces will have to be *unequally* spaced. The equal spacing of texture is a highly probable property of environmental surfaces, and this makes the even-textured regular tunnel still more highly probable. Nevertheless, it is not specified in the projection.

A fixed single-point projection does not specify any particular member of its family of transformations, but it does specify the family. In our example a family of cone-like surfaces is required whereas a family of pyramid-like surfaces is excluded. One member of the family is the "picture" of a regular cylinder, but the "picture" corresponding to Fig. 3 is *not* a member. The ambiguity is only partial.

If, however, instead of the "frozen" projection of Fig. 1, we consider a dual projection to two points or a fluid projection to a moving point, the member of the family of surfaces becomes geometrically determined and the ambiguity disappears. The "information" is supplied by the difference between different projections to different points in space. Consider for instance the noncongruence or disparity shown in the two parts of Fig. 2. This difference is unique. It is obtainable only from an evenly striped cylinder, not from a concave and unevenly striped cone, not from perspective rings on a plane surface, and not from a convex unevenly striped cone. The

disparity or deformation of the two projections, whether simultaneous or successive, is in correspondence with the depth of the object from which they issue.

A psychological hypothesis is now possible. If we assume that human *Os* can respond to such disparities, it might be predicted that, whereas the ambiguous stimulus may yield a variable percept, a dual or fluid projection will yield an invariable percept. In other words, with the use of two eyes, or with normal head movement, or both, the tridimensional shape of an object should become compellingly visible where it was not so previously.

There is reason to believe that a fixed single-point projection can induce variable perceptions. A photograph is a physical record of a fixed single-point projection, and it is possible to assert both that one perceives depth in a photograph and that one perceives it as flat—assertions which seem to be contradictory. The monocular view of an ordinary room from one position, such as a peephole, is a fixed single-point projection; accordingly a physical room may be constructed which is radically distorted from the ordinary room (if the proper transformation of rectangles into trapezoids is made) and yet appear ordinary through the peephole. But this perception may change to that of a distorted room if only *O*'s attitude or expectation is altered (**5**). Anyone who holds his head still and closes one eye gets a fixed single-point projection; one can then usually observe that the scene before him loses some degree or quality of its depth, and many psychologists have described this experience as a flat patchwork of color "sensations." However, most *Os* have great difficulty in seeing it as literally flat and the impression fluctuates. Among the controversies aroused by the theory of depthless sensations, one fact is clear: the kind of perception obtained with fixed single-point projection depends on the attitude of *O*. The optical pseudo tunnel provides a good opportunity to study perception under these conditions.

The Zero Gradient of Texture Density

Consider the pseudo tunnel shown in Fig. 3. As a fixed monocular stimulus it has a zero gradient of density of optical texture. It should therefore specify a family of objects the most probable of which is a flat frontal surface with a texture of evenly spaced rings. Tridimensional members of the family are possible but they have to be surfaces with *uneven* textures. A regular tunnel with even texture is excluded. The question is whether it will be *seen* as a flat frontal surface or, rather, how invariably it will be so seen. The senior writer has implied in earlier publications that a gradient of texture density is by itself a stimulus for the phenomenal slant of a surface (**2**, Chap. 6) but the present hypothesis is that it is by itself only an *ambiguous* stimulus.

In order to answer this question, two experiments were performed with an optical tunnel of the type shown in Fig. 3, that is, with accelerated spacing of apertures so as to provide optical rings of equal visual angle. In the first experiment 8 sheets were used and in the second 20. In the second, moreover, the nearest aperture was placed farther from the eye (176 cm.) so that the stimulus array was both smaller and denser. In both arrangements the farthest aperture was more than 6 m. from the eye. In the first arrangement the density was insufficient to yield a high degree of surface quality, but in the second it was. Observations were made with one eye occluded and with the head fixed by a biting-board. Naive *Os* were given 5-sec. exposures under these instructions: "Look carefully at what is behind the window. It is nothing complicated or surprising. I will ask you to describe it later, that is, to tell me how you would go about making something like it. Remember to keep your head motionless."

The 17 *Os* of the first experiment gave reports which were then classified. But they did not fall into categories indicating clearly either a bidimensional or a tridimensional perception. Most of the reports (59%) were of something ambiguous,

or of a thing which fluctuated between flat and deep. Of the remainder, about half fell into each category.

The 20 *O*s of the second experiment had much stronger impressions of a surface, and it was easier to classify the reports with respect to the flatness of the percept. The reports included answers to questions about the continuity of the surface (if one was seen) and about the amount of depth (if any was reported). For 25% of the *O*s the surface seen was perfectly flat and for 75% it had some depth. Examples of the latter were a "striped vase," a "funnel," or a "Mexican straw hat" (convex at the center and concave at the periphery). The depth perceived was always shallow and in no case did it even approach the actual depth of the tunnel, which was over 18 ft. The mean estimate of the depth of the object seen, omitting the flat cases, was 6.4 in. That the surface appeared continuous is indicated by the fact that only one *O* reported seeing anything like edges.

The conclusion must be that a peripheral-to-central gradient of texture density (in this case zero) does not by itself compel a perception of corresponding slant (in this case flat). This suggests that gradients of density in general are not in isolation determining stimuli for impressions of slant, as the geometrical considerations would predict. The psychophysical correspondence is not perfect. Nevertheless, as will appear later, they appear to determine *something* in the perceptual process. Perhaps it is a relationship between apparent slant and apparent spacing of texture elements. It would seem that the senior author has overstated the case in the past for the definiteness of fixed monocular depth perception.

The fact that a bidimensional percept can be the outcome of monocular stimulation from an 18-ft. tridimensional source implies strongly that the supposed cue of accommodation is not very significant for depth perception at these distances.

The Effect of Adding a Positive Parallactic Gradient to the Zero Density Gradient

If fixed monocular stimulation is not compelling for depth, how about fluid binocular stimulation, either or both? In the latter of the two experiments described above, all *O*s were given another 5 sec. presentation using both eyes and a third presentation using both eyes without the biting-board. The instructions were the same and the 20 reports were treated as before but they all fell into a new category. The surface seen had extended depth approaching the actual depth of the setup. Estimates were made in terms of feet, not inches, and the mean estimate was 5.8 ft. for binocular fixed vision and 6.2 ft. for binocular vision with head movement. The difference is not significant. So great was the contrast between the thing seen on first presentation and the thing seen on later presentations that a few *O*s expressed the suspicion that *E* had "done something" to the apparatus. In effect the object became a tunnel instead of something shallow or flat when parallactic stimulation was added. According to the reports of trained *O*s, as distinguished from naive *O*s generally employed in these experiments, either binocular or movement parallax would convert the object into a tunnel, and both together were not noticeably more effective than either alone.

The trained *O*s could mostly describe an unusual feature of the tunnel surface under these conditions: that its stripes or markings were not

evenly spaced. They seemed to become wider as the surface receded.

The conclusion must be that when a positive gradient of disparity or deformation is added to a zero gradient of density, from periphery to center of the stimulus array, the qualities of slant and recession appear in the perception. They determine these qualities despite the gradient of density, which seems to be wholly ineffective in this situation! There is some evidence that they do so only at the cost of deforming the phenomenal texture of the surface, i.e., making its elements unequally spaced.

The Effect of Adding a Negative Parallactic Gradient to the Zero Density Gradient by Means of a Pseudoscope

A check on the results of the last experiment can be obtained by introducing an "artificial" gradient of disparity to the stimulus instead of the "natural" gradient produced by using two eyes. If this peripheral-to-central gradient is reversed in direction, it would have to be predicted that the qualities of slant and recession would be reversed in the percept and, accordingly, the interior surface of a tunnel would be converted into the exterior surface of a truncated cone. Just such a reversal of the gradient can be produced by a pseudoscope (1, Ch. 7).

The effect of a pseudoscope can best be understood by turning to Fig. 2 and considering the horizontal disparities of the component margins in the two pictures. If each were separately inverted (rotated 180°), the effect would be that obtained with a lenticular pseudoscope, which is a pair of inverting lens systems. It may be noted that the disparities will then have been reversed in direction, which is to say that the gradient of disparity has been reversed. Precisely the same reversal will be produced by overturning each picture right for left out of the plane of the paper. This effect is obtained with a prism pseudoscope, which is a pair of totally reflecting prisms. These pictures, it should be remembered, represent the perspective of a "normal" optical tunnel with equally spaced apertures.

An ordinary scene like a room when observed through a lenticular pseudoscope looks upside down, and when observed through a prism pseudoscope looks reversed from right to left. Anomalies of depth may appear but the surfaces of the room do not appear to have been turned inside out. A better scene for pseudoscope viewing is a symmetrical object like a tunnel, which keeps the same bidimensional form when inverted or reversed. The ideal object for the purpose is probably the special optical tunnel used in the foregoing experiments which has the perspective of a flat surface.

A lens pseudoscope was mounted in front of the tunnel with its tubes parallel to and equidistant from the tunnel axis. As a control for the main experiment a small group of naive *O*s, in complete ignorance of the setup, were asked to apply an eye to one tube of the instrument and to report "what was there." The reports were comparable to those of the first experiment with monocular vision; the object seen was a surface of concentric rings, either flat or with shallow depth, the depth being for some *O*s concave, for others convex. The substitution of a lens system for free viewing had little effect on visible depth. Another group of 18 naive *O*s were asked to apply *both* eyes to the instrument and report. All without exception saw a striped convex truncated cone, or the equivalent, protruding from a background (presumably the nearest plastic sheet). The impression was immediate for nearly all *O*s, and the phenomenal object was stable. The prediction about reversed gradients of disparity and the slant of a surface is therefore confirmed.

Observers with knowledge of the setup see the same thing as *O*s without knowledge. Observations with a prism pseudoscope give the

same result. The convex phenomenal surface seen in this situation, as compared with the concave phenomenal surface seen with binocular vision of the tunnel, is nearer, smaller, and tapered instead of cylindrical. Its apparent distance, size, and shape are probably such as would be expected for a projective transformation of the interior surface of a cylinder into the exterior surface of a truncated cone—as if the cylinder had been pulled inside out within the defining sheaf of light rays.

When a pseudoscope is used with a "normal" optical tunnel (Fig. 1 and 2) instead of the special tunnel with the zero gradient of density, the outcome is not the same as above. In this situation a gradient of texture density is in full opposition to the gradient of disparity. Many but not all *O*s can see the protruding truncated cone, nearer and smaller than the receding cylinder, but the percept is unstable and gives way at times to the cylinder or to some compromise between them. When the cone is seen, it has a very odd distribution of stripes, extremely small and close together at the front but very large and far apart at the rear. This is further evidence to show that *when a gradient of disparity wins out over a conflicting gradient of texture density it does so at the cost of giving the phenomenal texture an uneven scatter or unequal spacing, that is, deforming it.*

When the cone is not seen but gives way to the cylinder or to a compromise, there are indications that *O* has lost his binocular fusion and is getting double imagery. The diplopia can usually be observed under these conditions. This suggests the hypothesis that *when a gradient of texture density wins out over a conflicting gradient of disparity it does so at the cost of producing double images.*

The implications of these results for ordinary visual perception.—The last three experiments have been stimulus situations where the gradient of density was isolated from or discrepant with the gradients of disparity and motion. In ordinary stimulus situations they are concomitant, not discrepant. In the everyday environment the reflecting elements of surfaces tend to be evenly spaced, both eyes are open, and the head moves. The implication is that the limited ambiguity of fixed monocular vision is removed when binocular or motion parallax is operative.

Summary

A method is described for inducing and controlling a perception of *surface* and *space*. Conclusions are: (*a*) There is evidence that surface quality depends on the density of transitions in the optical stimulus. When the transitions are absent, surface quality disappears. (*b*) There is evidence that the gradient of texture density in isolation need not always determine the qualities of slant and recession. It is ambiguous, but only with respect to the members of a family of surfaces. (*c*) There is evidence that the gradients of textural disparity and motion can determine the qualities of slant and recession. Presumably when the gradients are concomitant, ambiguity is removed.

References

1. Carr, H. A. *An introduction to space perception.* New York: Longmans, Green, 1935.
2. Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1951.
3. Gibson, J. J., & Cornsweet, J. The perceived slant of visual surfaces—optical and geographical. *J. exp. Psychol.*, 1952, **44**, 11–15.
4. Gibson, J. J., & Waddell, D. Homogeneous retinal stimulation and visual perception. *Amer. J. Psychol.*, 1952, **45**, 263–270.
5. Ittelson, W. H. *The Ames demonstrations in perception.* Princeton: Princeton Univ. Press, 1952.

(Received July 28, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

THE ROLE OF SIMULTANEOUS CONTRAST IN BRIGHTNESS CONSTANCY [1]

H. LEIBOWITZ, NANCY A. MYERS, AND P. CHINETTI

University of Wisconsin

Visual perception may be considered as a compromise between predictions made on the basis of retinal image theory and "constancy" responses which tend to preserve critical perceptual characteristics under various viewing conditions. Color, brightness, or whiteness constancy refer to the observation that we perceive objects as tending to maintain a characteristic color or whiteness despite variations in the quality or quantity of luminous flux in the retinal image. A number of variables have been considered as factors relevant to the elicitation of constancy responses (**3, 4, 5, 6, 8**). Among these the role of simultaneous brightness contrast or induction has been considered, although the experimental studies have considered only relatively narrow ranges of variables.

The purpose of the present experiment is twofold. (*a*) We describe the functional relation between the illuminance incident on a test object and the luminance of a photometric field with which it is matched. (The "law of brightness constancy" would predict that the matched luminance would remain invariant; the "additivity law of luminance," that the function would be linear with a slope of unity.) (*b*) We determine the effect of simultaneous brightness contrast as a parameter of this function.

[1] This research was supported with funds provided by the Wisconsin State Legislature. The authors wish to express their appreciation to Drs. W. J. Brogden and D. A. Grant for valuable assistance in the preparation of the manuscript.

In order to investigate contrast as an experimental variable, it is necessary to define the conditions under which it occurs. Recent experiments, designed to provide the necessary specifications, indicate that these include at least two visual stimuli with different luminance values which coexist in the visual field (**2, 7**). Contrast is defined experimentally in one direction only: as the inhibition of the excitation produced by the stimulus of lower luminance.

Apparatus and Procedure

A room 8 ft. long, 8 ft. high, and 4 ft. wide was constructed of masonite and wood for the experiment. A top schematic view is presented in Fig. 1A. The room is divided into two chambers for half its length, the left chamber containing the standard stimulus (SS), made from the No. 9 Munsell gray (reflectance 19.77%), which is mounted in the center of a background 18 in. wide by 24 in. high. Variable illuminance for this side of the apparatus is provided by a series of reflector flood lamps located at various distances from the stimuli and utilized in combinations appropriate to provide a range of variation in illuminance of more than one million to one. The lights are hidden from *S*'s view and are operating on a 60-cycle, 110-v. circuit. Luminance ripple, determined with a photocell and oscilloscope, is less than 5%.

The right chamber was dark except for an opening (CO) through which the comparison stimulus, one of the series of 17 Munsell grays, was viewed. This opening is the same size and is located in a position comparable to the standard stimulus. The series of comparison stimuli are mounted in order on a wheel (CW) which can be rotated by means of a 1-rpm reversible motor (M) so that any desired stimulus is visible. Illuminance for the comparison field is provided by a reflector flood located outside of the room. Since the reflectance ratio of the lightest to the darkest gray is less than 100 to 1, neutral tint filters (F) were

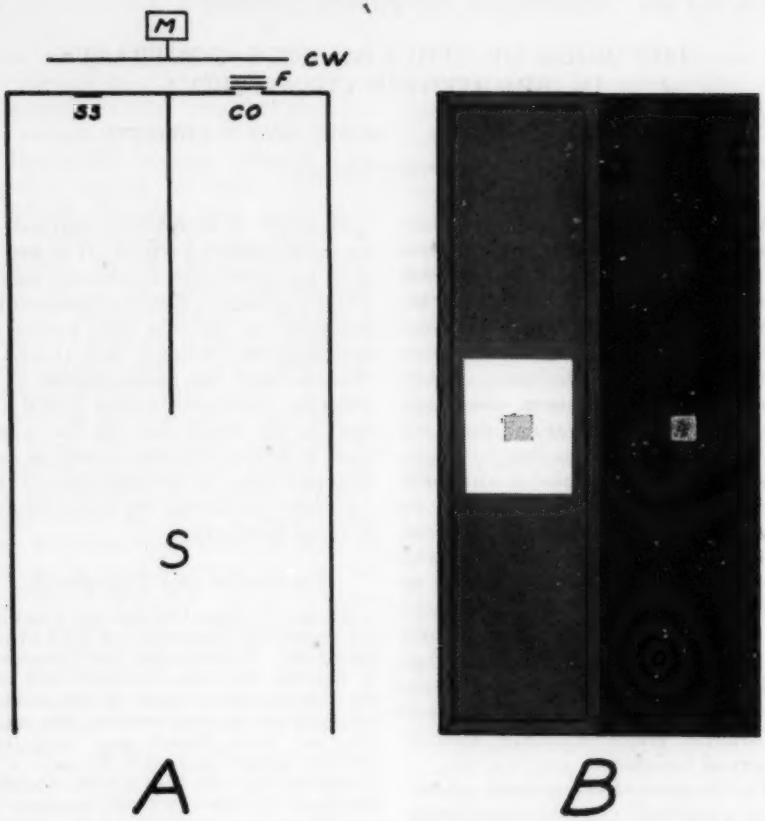

FIG. 1. (A) Schematic diagram of the top view of the apparatus; (B) Illustration of *S*'s view of the apparatus.

placed behind the comparison stimulus opening to provide the proper luminance range for matching with the standard stimulus.

The *S* was seated in front of the partition separating the two chambers so that he could view the standard or comparison fields by a slight movement of the head. Both square fields are 6 ft. from *S*'s eyes, and each subtends a visual angle of 1°. The operation of the switch which controls the motor rotating the comparison wheel is explained to *S* and instructions are given to "choose the comparison stimulus which looks the most like the standard." Except for the comparison, standard, and background fields, all parts of the apparatus are painted a dull black (reflectance 1.78%). The *S*'s view of the apparatus is illustrated in Fig. 1B.

Thirty *S*s from elementary psychology classes, who were unaware of the purpose of the experiment, served for one session each. They were assigned randomly to three groups of 10 *S*s each. All *S*s matched the standard stimulus with the comparison at the following seven levels of illuminance: .00073, .0046, .012, .51, 50.1, and 631, and 1,580 ft.-candles.[2] The procedure for the three groups differed only with respect to the background against which the standard was viewed. For Group 1 the background was black flocked paper (reflectance .85%), for Group 2 white cardboard was used (reflectance 62%), while Group 3 observed against a gray of the same reflectance as the standard. Half of the *S*s in each group were presented the illuminance levels in increasing, the other half in decreasing order. Six judgments utilizing the method of adjustment were made by each *S* at every illuminance level. Between judgments the illuminance on the

[2] Luminance values in foot-Lamberts may be calculated as the product of the illuminance in foot-candles and the reflectance of the surface in question.

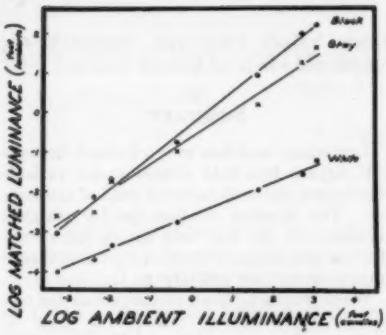

FIG. 2. Log matched luminance of the comparison field as a function of the log ambient illuminance on the test stimuli for three background conditions. On this plot a slope of unity indicates agreement with the "additivity law of luminance." A slope of zero represents the "law of brightness constancy." The "white" background has a reflectance of 62.00%, the "gray" background 19.77%, and the "black" background .85%.

comparison field was extinguished and *E* offset the comparison wheel. No time limit was placed on *S*'s responses. A single session, consisting of instructions, dark adaptation, and 42 judgments, took about 1 hr.

RESULTS

The mean log luminance of the comparison stimuli required to match the standard stimulus at seven levels of log illuminance and for three background conditions are presented graphically in Fig. 2.[3] Each datum is the mean of the average matched luminance values for 10 *S*s. On this plot a line with a slope of one would represent a prediction in terms of the "additivity law of luminance," such that an increase in the ambient illuminance produces the same increase in the matched comparison field luminance. It predicts a function such as would be obtained with a photosensitive recording device based solely on energy relationships. At the other theoretical extreme, the law of brightness constancy, which predicts that matched luminance should remain invariant at all levels of ambient illuminance, would be represented here by a line of zero slope.

The slopes of the lines fitted to the data by the method of averages are .82 for the "black," .67 for the "gray," and .42 for the "white" backgrounds. With the exception of the point for the lowest illuminance level for the "gray" background, all three sets of data are closely described by straight line functions on the double log plot. The data for the "black" and "gray" background conditions differ from the "white" background not only in slope, but also with respect to their position on the ordinate axis, the latter falling from one to three log units lower on the log matched luminance axis. An Alexander trend analysis (1), summarized in Table 1, shows the differences between group slopes and group means to be significant at the .001 level of confidence.

DISCUSSION

The functions relating the matched luminance of a test field and the ambient illuminance are satisfactorily described by straight line functions on a double log plot. Similar results have been previously reported by Hsia for a more restricted illuminance range (5). All of the functions, irrespective of background

TABLE 1

ANALYSIS OF TRENDS OF LUMINANCE MATCHES UNDER THREE BACKGROUND CONDITIONS

Source of Variation	*df*	Mean Square	*F*
1. Between group means	2	216.67	92.95**
2. Between group slopes	2	30.73	151.38**
3. Between individual means	27	31.47	11.48**
Error estimate for (1)			
4. Between individual slopes	27	9.62	3.51**
Error estimate for (2)			
5. Group deviations from estimation	19	2.02	1.05
6. Over-all slope	1	485.98	4787.98**
7. Over-all deviations from linearity	5	1.54	3.03*
8. Individual deviations from estimation	135	13.71	
Error variance estimate			

* .05, ** .001.

[3] Individual data from which Fig. 2 was constructed have been deposited with the American Documentation Institute. Order document No. 4517, remitting $1.25 for microfilm or $1.25 for photocopies.

conditions, exhibit slopes which are between predictions based on the theoretical "additive" and "constancy" laws. Such results are typical in visual perception, the data representing a "compromise" between retinal image theory and constancy.

With respect to the variable of simultaneous contrast, it is significant that the tendency towards constancy is minimal with the "black" and "gray" backgrounds. The luminance of these backgrounds is less than the test field and, as would be expected from contrast studies (**2, 7**), they do not inhibit the excitation produced by the test field. This absence of simultaneous contrast with the "black" and "gray" backgrounds is indicated by the observation that the luminance matches for these conditions do not differ greatly from predictions made on the basis of energy relationships. However, for the "white" background of higher luminance than the test field, both simultaneous contrast and the tendency towards constancy are observed. Contrast is manifested by the inhibition of the excitation produced by the test field and results in the shift of the "white" function to lower matched luminance values. Constancy is reflected by the lowering of the slope of this function, in relation to the slopes for the "black" and "gray" background conditions, to a value which is about halfway between the additivity and constancy laws. Thus, the tendency towards constancy and the presence of simultaneous contrast occurred under the same stimulus conditions.

Constancy behavior is considered to depend upon the co-existence in the visual field of stimuli other than *the* discriminative stimulus (**4**). The present study indicates that a background of higher luminance than the test field is an important *one* of these stimuli, operating to produce the tendency towards constancy over a wide range of ambient illuminances. The role of other variables, such as familiarity, instructions, exposure duration, personality factors, etc., remains to be clarified for experimental conditions in which the illuminance values employed approach the functional range of human vision.

Summary

Luminance matches were obtained between a 1° square test field viewed under variable illuminance and a photometric field of the same size. The function relating the log matched luminance of the test field to the log of the ambient illuminance is linear over an illuminance range of more than 1,000,000 to 1.

When the background against which the test field is viewed is either "black" or a "gray" of the same reflectance as the test field, the matches do not differ greatly from predictions based on energy relationships.

When the test field background is a "white" of higher reflectance than the test field, the slope of the function is less and the data are shifted to a lower position on the log matched luminance axis. The change in slope is a manifestation of brightness constancy. The shift in the ordinate position is a result of simultaneous contrast.

It is concluded that one of the cues which tends to produce brightness constancy is a background of higher luminance than the test field.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Diamond, A. L. Foveal simultaneous brightness contrast as a function of inducing and test-field luminances. *J. exp. Psychol.*, 1953, **45**, 304–314.
3. Feldman, S., & Weld, H. P. Perceiving. In E. G. Boring, H. S. Langfeld, and H. P. Weld (Eds.), *Psychology*. New York: Wiley, 1935. Pp. 274–299.
4. Graham, C. H. Visual perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Chap. 23.
5. Hsia, Y. Whiteness constancy as a function of differences in illumination. *Arch. Psychol.*, N. Y., 1943, No. 284.
6. Katz, D. The world of color. (Trans. by R. B. Macleod and C. W. Fox.) London: Paul, 1935.
7. Leibowitz, H., Mote, F. A., & Thurlow, W. R. Simultaneous contrast as a function of separation between test and inducing fields. *J. exp. Psychol.*, 1953, **46**, 453–456.
8. Woodworth, R. S., & Schlosberg, H. *Experimental psychology*. (Rev. Ed.) New York: Holt, 1954.

(Received August 30, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

EFFECT OF PHOTOMETRIC BRIGHTNESS ON JUDGMENTS OF DISTANCE [1]

JOHN COULES

Boston University[2]

Distance judgments of an object are partially a function of the actual distance of an object from the eyes. A variety of studies have shown that brightness (**1**), size (**2**), color (**10**), contrast (**4**), monocular viewing (**3**), and convergence (**6**) influence the judgments of distance. Investigation of any one of these variables requires that the remaining ones be held constant or controlled. On some occasions previous investigators of the effect of brightness have failed to control such relevant factors as lateral eye movements and color temperature. In spite of this criticism it has been shown that under binocular conditions brighter objects appear nearer (**2, 7, 8**).

Helmholtz explains this phenomenon in terms of irradiation (**7**). Light from the object would irradiate as it passes through the ocular media and is projected upon the retina. If the theory of irradiation explains the phenomenon, then it follows that *S*s, under monocular viewing conditions, would judge brighter objects nearer. It is necessary to determine whether this is the case.

Psychophysicists are concerned with the problem of equivalence, that is, finding alternative stimulus complexes for which a given attribute of responses remains constant (**9**). In reference to the brightness variable and the actual distance of an object the following question is in order. Would a brighter object farther away be judged equal to a dimmer object nearer to *S*? Graham states that stimulus-stimulus functions would indicate this equivalence (**5**).

Two experiments were performed. The purposes of Exp. I were (*a*) to determine the functional relationship between photometric brightness and judgments of distance under binocular and monocular viewing conditions,[3] and (*b*) to obtain stimulus-stimulus functions for brightness and object distance. The purpose of Exp. II was to determine whether the effect of brightness on distance judgments could be generalized by varying the absolute level of brightness.

Apparatus

The apparatus consisted of two light sources and *S*'s station.[4] The *S*'s head was firmly positioned in a modified welder's mask through which he could observe the two light sources. The light source on *S*'s left was fixed and was the standard light stimulus (SLS). The light source on *S*'s right was spatially variable and was the variable light stimulus (VLS). Each stimulus was an evenly illuminated opal glass disc.

[1] This paper is abstracted from a dissertation submitted to the Graduate School of Boston University in partial fulfillment of the requirements of the Doctor of Philosophy degree. The research was carried out in the Psychology Section of Boston University, Optical Research Laboratory under contract with the Photographic Reconnaissance Laboratory, Wright Air Development Center, USAF (Contract AF 33-038-15615). Thanks are due to Dr. L. J. Reyna, who directed this research. Valuable assistance was received from Professor J. M. Harrison.

[2] Now at the University of New Hampshire.

[3] Photometric brightness as defined by the American Standards Association. Illuminating engineering nomenclature and photometric standards, 1942.

[4] The apparatus was designed by E. J. Robinson and the author for a program of research in visual discrimination.

Each light source consisted of a wooden housing unit and a projecting tubular baffle system. The SLS was fixed at 25 ft. from *S*'s station in order to exclude accommodation as a relevant variable. The VLS was mounted on a 10-ft. long steel track, the mid-point of which was also 25 ft. from *S*. Thus *E* could place VLS at any position on the track at a maximum distance of 5 ft. in either direction from the mid-point. In order to maintain a constant distance between the two stimuli, the track for VLS was arranged so that the light sources always converged at an angle of 6°4′ upon a mid-point between the eyes of *S*.

A baffle tube was added in front of each light source to eliminate stray light from illuminating the test room which might provide *S* with extraneous cues. Five circular aluminum baffles, with a 2.9-in. diameter opening in each, were placed at 5-in. intervals through the length of the 25-in. projecting tube. The entire apparatus was painted a flat black.

Within each housing unit a 6-v. GE medical lamp was mounted. To obtain a homogeneous disc of light these lamps were mounted on a rotatable supporting base plate. When the base plate was moved to the front end of the unit, the center of the lamp's filament was 3 in. from the circular opal glass. The opal glass was mounted adjacent to the housing at its junction with the baffle tube. The white opal glass, 3.5 in. in diameter, provided a surface that was apparently uniform even at close range.

Between the housing and the opal glass disc an aluminum knife edge was installed to eliminate light scatter. The knife edge on SLS was 2 in. in diameter, while that on VLS was 3 in. This difference was necessary because a diaphragm in front of the opal glass on VLS had to be opened wider than the 2-in. diameter of SLS.

The diameter of the diaphragm varied with the movement of VLS so that the visual angle remained constant as VLS was moved along the track. The diaphragm was continuously varied from 1.6 to 2.4 in. by means of a system of levers. One of the levers was connected to a piano wire mounted on the right-hand side of the frame of VLS. In addition, this arrangement kept the illumination at the eye relatively constant as VLS was moved a distance of 10 ft. along the track. The measured change in illumination, with a photoelectric cell, was not more than 3% (averaged over the 10-ft. travel), which is a negligible amount.

Two operations were used to obtain changes in illumination on the opal glass without introducing color temperature changes: (*a*) moving the lamp in the housing unit away from the opal glass, and (*b*) placing filters in front of the glass disc. The *E* could vary the position of the lamp by means of a calibrated rod which was connected to the base of the lamp. The change in illumination was shown to follow an inverse-square law. Provision was made to install filters at the outer end of the baffle tubes. The filters were of the Wratten neutral-density type which transmitted 37% and 9.5% of the light from the opal glass.

The electrical system (100-v. AC) consisted of a voltage regulator, a variable transformer (Variac) and a 6-amp. filament transformer. The two lamps were in parallel from the last unit in the series.

Procedure

Preliminary investigation showed: (*a*) an average jnd distance value of 4 in. was necessary to insure differential distance responses, and (*b*) brightness ratios of 7.5:1 (VLS:SLS) and less had to be used to obtain differential responses of nearer and farther. The ratios chosen fell within these limits and were: 6.25:1, 2.5:1, 1:1, 1:2.5, and 1:6.25. At any one ratio, except the 1:1, the difference between the two lights was well above the difference threshold for brightness discrimination.

Experiment I.—The objective of this experiment was to determine the functional relationships between ratios of brightness and judgments of distance under binocular and monocular conditions. Two *S*s with uncorrected visual acuity of 20/20 (Snellen) in each eye were employed. B. M., a 20-yr. female, was a senior psychology major. R. B., a 25-yr. male, was a first-year graduate psychology student.[5]

The *S*s participated in 20 daily sessions over a period of four weeks. The *S* was told that his task was to state whether the light source on his left or the one on his right was nearer. The method of constant stimuli was used to present the 64 stimulus combinations. The order of stimulus presentation was randomized by means of a table of random numbers for brightness ratios and distance intervals. Each *S* was tested first under binocular conditions (40 runs per stimulus presentation) and then under monocular conditions (15 runs). Under the monocular condition *S* used his dominant eye. Subject B. M. used her left eye and R. B. his right eye.

In this experiment only one brightness level was used (13.0 ft.-l.). In each stimulus presentation one of the light sources had a brightness of 13.0 ft.-l. and the other light source was decreased to yield the five brightness ratios.

[5] Only the data for Subject R. B. are reported. Similar results were obtained for Subject B. M.

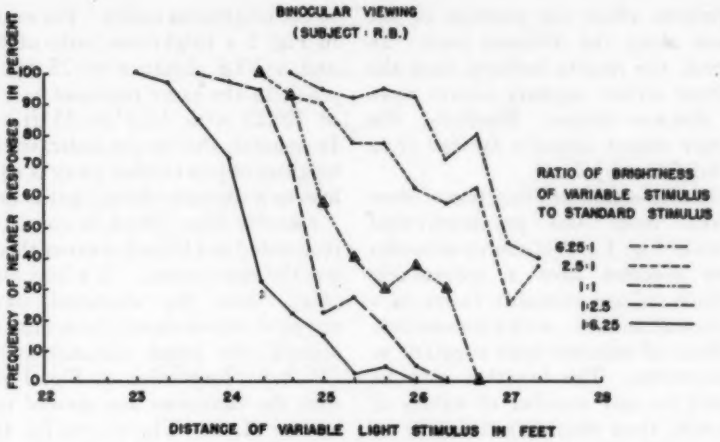

FIG. 1. Percentage frequency of "nearer" responses as a function of the distance of the variable light stimulus (VLS) and the brightness ratios under binocular conditions: Subject R. B.

All the data were plotted in terms of the "nearer" responses to the VLS ("right" response).

Experiment II.—This experiment was performed to determine whether the relationship between brightness and distance judgment found in the first study would hold when the absolute level of brightness was varied.

Two female *Ss* with uncorrected visual acuity of 20/20 in each eye were used. Subject J. D., age 20, had a right-eye dominance and M. M., age 21, had a left-eye dominance.

In this study three levels of brightness were used.[6] The instructions, apparatus, experimental room, distance intervals, and brightness ratios were the same as in Exp. I. All determinations were made under binocular conditions.

The up-down method of stimulus presentation was used to obtain the various points of subjective equality.[7] This procedure consisted in presenting *S* with the light sources set at a particular brightness ratio. The *Ss* were first run at the 13.0 ft.-l. level with the five ratios in the order 6.25:1, 2.5:1, 1:1, 1:2.5, and 1:6.25. This procedure was repeated for the 3.33 and .035 ft.-l. levels. At the beginning of each session VLS was placed at the 27-ft. distance position of the track and *S* was required to state whether the left or the right light was nearer. If *S* said "left," then VLS was brought 4 in. nearer. If *S* made the same response on the next trial, then VLS was brought 4 in. nearer. This procedure was continued until *S* reversed his response. After the first reversal, *E* always set VLS one increment in the opposite direction to *S*'s last response. The entire length of the track (10 ft.) was used to produce a total of 31 distance positions 4 in. apart. Sixty responses were obtained for each brightness ratio and level.

[6] When the brightness level was increased above 13.0 ft.-l., a noticeable "halo" effect occurred around the light source.

[7] Dr. G. A. Miller, personal communication, 1953.

RESULTS

Experiment I

Binocular viewing.—Psychophysical curves (Fig. 1) were drawn to show the effect of brightness on distance judgments. The function for the 1.1 brightness ratio may serve as a reference curve. When VLS is brighter than SLS by ratios of 2.5:1 and 6.25:1, the corresponding curves at these ratios are displaced to the right of the reference curve, indicating that VLS distance settings greater than 25 ft. are being responded to as "nearer." When VLS is dimmer than SLS by ratios of 1:2.5 and 1:6.25, the corresponding curves are displaced to the left of the reference curve. Therefore, systematic changes in

brightness affect the position of the curves along the distance scale. In general, the results indicate that the brighter object appears nearer than the dimmer object. Similarly, the dimmer object appears farther than the brighter object.

Stimulus-stimulus functions were derived from the psychophysical curves in Fig. 1. A stimulus-stimulus curve specifies how a measurable attribute of one stimulus varies as a function of another, with a measurable attribute of response kept constant as a parameter. The function may be plotted for any number of values of response, thus obtaining a family of curves. The distance value that corresponds to the 50% frequency of "nearer" responses may be determined from each curve in Fig. 1. These values are now plotted against the brightness ratios. Figure 2 illustrates the functions for the 25%, 50%, and 75% values of response frequency. Inspection of the curves shows that the distance of VLS is an increasing function of the logarithm of the brightness ratios. For example, in Fig. 2 a brightness ratio of 6.25:1 and a VLS distance of 25 ft. 2 in. result in the same response as a ratio of 1:6.25 with VLS at .23 ft. 4 in. In general, the results indicate that a brighter object farther away is equivalent to a dimmer object that is nearer.

Observer bias.—Both *S*s consistently responded in a biased manner throughout the experiment. If a bias did not exist, then the expected psychophysical curves should be symmetrical around the equal distance position (25 ft.). Inspection of Fig. 1 shows that the functions are skewed to the left for R. B. The curves for B. M. (not shown) were skewed to the right. In addition, when the constant errors were determined, R. B. had an error of + 2 in. and B. M. had an error of − 4 in. Thus, for R. B. the VLS is equivalent in distance to the SLS when the VLS is farther away. The difference in ocular dominance for these *S*s may account for the observed differences. R. B., who had a right-eye dominance, made more "nearer"

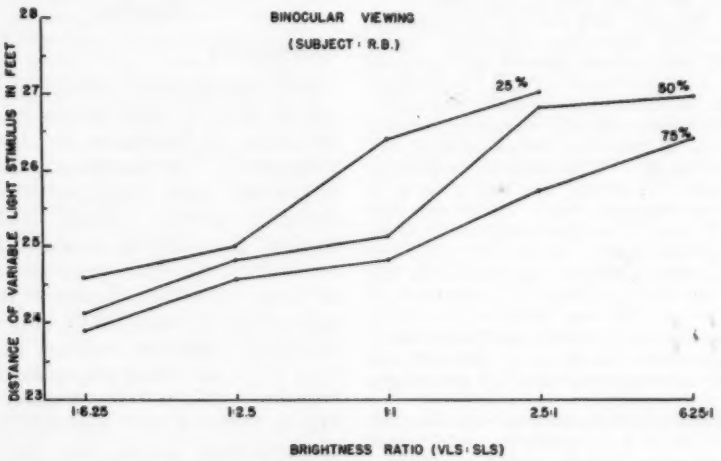

FIG. 2. Distance of the variable light stimulus (VLS) as a function of the logarithm of the brightness ratios with response frequency as a parameter: Subject R. B.

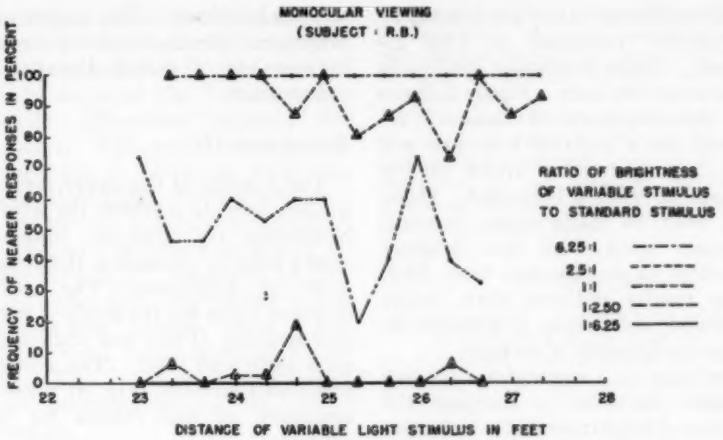

FIG. 3. Percentage frequency of "nearer" responses as a function of the distance of the variable light stimulus (VLS) and the brightness ratios under monocular conditions: Subject R. B.

responses to the light stimulus on his right although it was farther away in many instances. Similarly, B. M., who had a left-eye dominance, made more "nearer" responses to the light stimulus on her left. This suggests that ocular dominance may play a role in the determination of distance judgments.

Monocular viewing.—The *S*s were able to judge the distance of VLS in the binocular part of the experiment when the brightness ratio was unity (Fig. 1). As VLS was brought nearer

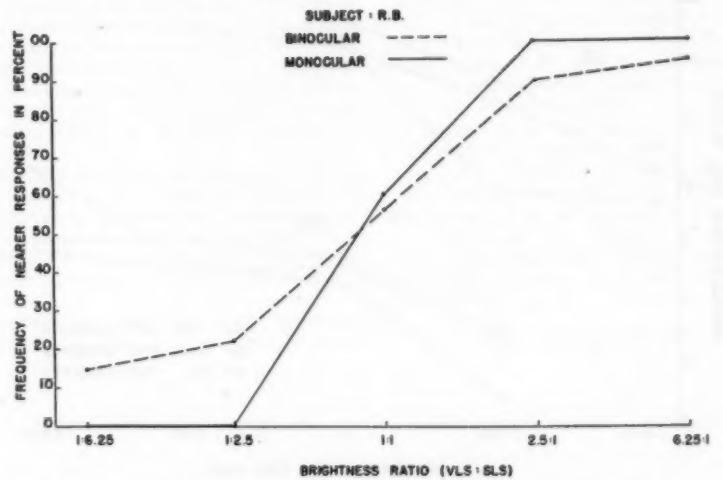

FIG. 4. Percentage frequency of "nearer" responses as a function of the logarithm of the brightness ratios: Subject R. B.

for all brightness ratios the frequency of "nearer" responses to VLS increased. Under monocular conditions this was not the case. Figure 3 shows that the frequency of "nearer" responses, for a ratio of 1:1, does not show any systematic trend as the distance of VLS is increased. However, both *S*s made more "nearer" responses when VLS was brighter regardless of its distance from SLS. These results indicate that, under monocular conditions, brightness influences judgments of distance.

Binocular vs. monocular, distance constant.—In order to compare the influence of brightness under binocular and monocular conditions, it is necessary to keep distance constant. Accordingly, the percentage frequency of "nearer" responses was plotted against the logarithm of the brightness ratios at the 25-ft. distance position from Fig. 1 and 3. Figure 4 illustrates this relationship. As the logarithm of the brightness ratio increases the percentage frequency of "nearer" responses increases. This suggests that brightness affects distance judgments independent of retinal disparity and convergence.

Experiment II

The purpose of this experiment was to demonstrate whether the effect of brightness on distance judgments would hold for changes in the absolute level of brightness. The average distance value for the point of subjective equality (PSE) was obtained for each ratio and level. The stimulus-stimulus functions (Fig. 5) present a summary of the results for both subjects.

Inspection of these curves at any brightness level shows that as the brightness ratio is increased there is an increase in PSE value. Thus, the effect of brightness on distance judgments operated on all three levels. In general, the PSE value for VLS is an increasing function of the logarithm of the brightness ratios. For both *S*s

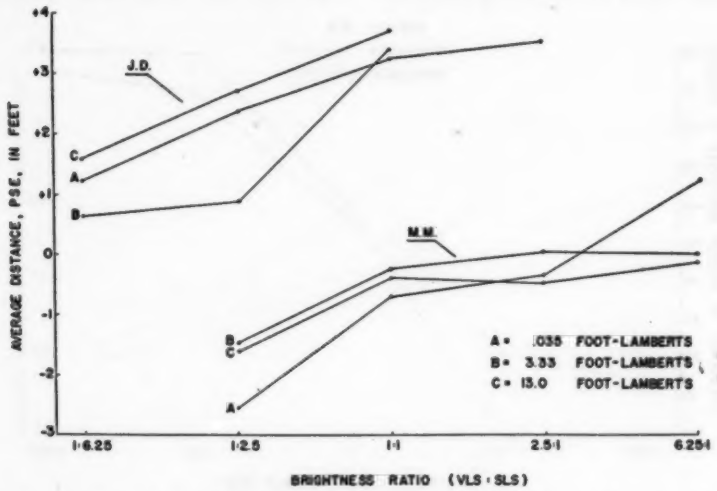

FIG. 5. Equal distance (PSE) in feet as a function of the absolute brightness levels and the logarithm of the brightness ratios.

there is no consistent increase or decrease in PSE as the brightness level (the parameter) is decreased.

Examination of Fig. 5 shows some emphatic differences between *S*s' judgments. PSE values could not be obtained for J. D. when VLS was brighter than SLS. For some brightness ratios this *S* reported that VLS was "nearer" although it was 5 ft. farther than SLS. Similarly, M. M., for one ratio (1:6.25), tended to say that VLS was "farther" when it was actually 5 ft. nearer than SLS. As found in Exp. I, the ocular dominance of *S*s corresponds to the above differences.

Summary

In Exp. I relative distance judgments, under binocular and monocular conditions, were obtained with two *S*s using the method of constant stimuli. The distance of the variable stimulus and the brightness ratio of the variable to the standard stimulus were varied. In Exp. II, two *S*s made distance judgments under binocular conditions, using the up-down method. In addition to the variables of Exp. I, the brightness level was varied.

An apparatus is described which permitted the presentation of two homogeneous discs of light whose brightness was varied. The variable light stimulus was mounted on a 10-ft. long track whose mid-point was 25 ft. from *S*. The standard stimulus was set at a distance of 25 ft. An attempt was made to control other relevant variables likely to influence judgments of distance.

In Exp. I psychophysical functions indicated that brightness is a factor in judgments of distance under binocular and monocular conditions. Stimulus-stimulus functions for binocular conditions showed that a brighter object farther away is equivalent to a dimmer object that is nearer.

In Exp. II the effect of brightness on distance judgments operated at two lower levels of brightness in addition to the level used in the first experiment. However, no consistent results were obtained between the brightness level and the point of subjective equality for a given brightness ratio.

The stimulus-stimulus functions for both experiments were similar. Observed differences between the subjects were attributed to ocular dominance. It was suggested that ocular dominance may account for the observer bias found in both experiments.

References

1. Ashley, M. L. Concerning the significance of intensity of light in visual estimates of distance. *Psychol. Rev.*, 1898, **5**, 595–615.
2. Carr, H. A. *An introduction to space perception.* New York: Longmans, Green, 1935.
3. Deyo, B. V. Monocular and binocular judgment of distance. *Amer. J. Ophthal.*, 1922, **5**, 343–347.
4. Fry, G. A., Bridgman, C. S., & Ellerbrock, V. J. The effect of atmospheric scattering on binocular depth perception. *Amer. J. Optom.*, 1949, **26**, 9–15.
5. Graham, C. H. Behavior, perception and the psychophysical methods. *Psychol. Rev.*, 1950, **57**, 108–120.
6. Grant, V. M. Accommodation and convergence in visual space perception. *J. exp. Psychol.*, 1942, **31**, 89–104.
7. Helmholtz, H. *Physiological optics.* (Trans. by J. P. C. Southall.) Vol. 2. Optical Society of America, 1924.
8. Ittelson, W. H. *The Ames demonstrations in perception.* Princeton, N. J.: Princeton Univ. Press, 1952.
9. Stevens, S. S. Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.
10. Taylor, I. L., & Sumner, F. C. Actual brightness and distance of individual colors when their apparent distance is held constant. *J. Psychol.*, 1945, **19**, 79–85.

(Received August 27, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

DEVELOPMENT OF RESPONSE GENERALIZATION GRADIENTS [1]

CARL P. DUNCAN

Northwestern University

When a correct response is strengthened by reinforcement, other responses which are ordered along some dimension of similarity to the correct response (generalized responses) are affected in a complex way. If these generalized responses occur as overt errors, they will, like all errors, decrease in occurrence. When considered as errors, all generalized responses are weakened, at least relatively, by reinforcement of the correct response. However, responses of different degrees of similarity to the correct response are not equally affected. Whatever the effect is on generalized responses of reinforcement for the correct response, it is known that the effect is proportional to the degree of similarity between the correct and the generalized response (**3, 4, 5**). In two of the papers just cited (**3, 4**) it was assumed that the effect was one of "parasitic reinforcement," by which was meant that when the correct response was reinforced, generalized responses were also strengthened to some extent. Further, it seemed reasonable to assume that the more similar a generalized response was to the correct response, the more it was strengthened by parasitic reinforcement. Thus there would be established a response generalization gradient, a gradient of strength of generalized responses which was directly proportional to the gradient of similarity between correct and generalized responses.

There is an alternative to the hypothesis that generalized responses are strengthened by parasitic reinforcement, and strengthened in direct proportion to the degree of similarity. The argument runs as follows. Since all responses except the correct one are, by definition, incorrect, generalized responses are errors. As such they must be suppressed or extinguished during learning. If we make the assumption that the extinction or relative weakening is greater for generalized responses which are less similar to the correct response, we have a process that would yield the response generalization gradient. In other words, instead of assuming that more similar generalized responses are more strengthened by parasitic reinforcement, we could assume that less similar generalized responses are more rapidly extinguished. This latter assumption may be denoted as "differential rate of extinction." Hull (**2**) has also suggested this idea, but with reference to stimulus generalization.

In the three papers (**3, 4, 5**) which have dealt with response generalization gradients it was not possible to contrast parasitic reinforcement with differential rate of extinction. In all cases the correct responses were meaningful adjectives, so it is

[1] This is a portion of a study performed under Air Force Contract No. 33 (038)-11396 between the Aero Medical Laboratory of the Wright Air Development Center, Air Research and Development Command, and Northwestern University. Thanks are due Dr. B. J. Underwood, Dr. V. E. Montgomery, Mr. Marshall Segall, and Mr. R. Dale Dick.

possible that a response generalization gradient for each correct response existed in some form prior to the beginning of practice. Further, the number of overt errors which could be considered to be generalized responses was not great even early in practice, and they rapidly decreased as practice proceeded. With such situations it is difficult to accomplish all of the following conditions. (*a*) There shall be available to *S* generalized (error) responses which differ from the correct response by definable degrees of similarity. (*b*) These generalized responses shall occur as overt errors in more than negligible numbers both early and late in practice. (*c*) Early in practice generalized responses of different degrees of similarity shall occur with approximately equal frequency, or at least with frequencies that do not correspond in any meaningful way with the gradient of similarity.

When these three conditions are met, it will be possible to determine if a response generalization gradient can be developed from scratch during original learning of a task, and to measure the gradient directly in terms of overt errors rather than in terms of a derived measure such as transfer (**3, 4**). Thus it should be possible to determine if response generalization gradients can be established by differential rate of extinction of generalized error responses, without resorting to the notion of parasitic reinforcement.

METHOD

Apparatus.—Since the apparatus has been described in detail elsewhere (**1**), it will merely be summarized here. The *S* learned a task consisting of six perceptual-motor paired associates in which the stimuli were colored lights flashed on a single ground-glass screen, and the responses were movements of a lever, held by the right hand, into six horizontal radial slots. The task was made difficult by simultaneously requiring *S* to hold another lever with the left hand. The left-hand lever essentially provided a steadiness test; *S* could not record correct responses with the right lever unless the left lever was being held, within narrow limits, in the vertical position.

The right-hand task is convenient for the purposes of this experiment. Since the responses are movements of a lever into horizontal radial slots, all responses are equally available for all stimuli. Errors are rather frequent initially and are not completely eliminated even after 180 practice trials (**1**). The spatial arrangement of the slots provides a situation such that for any correct slot the other five slots are arranged on a bidirectional gradient of spatial similarity around the correct slot. Thus the task at least meets those conditions, specified earlier, which have to do with the nature of the task; i.e., that there shall be error responses which fall along a gradient of similarity to each correct response, and that these generalized responses shall occur in sufficient numbers as overt errors both early and late in learning.

The apparatus was set to provide, automatically, 20-sec. trials with 10-sec. intertrial rests, and within a trial was subject paced, i.e., a new stimulus flashed on the screen immediately upon *S*'s recording a correct response.

Since it was necessary for this experiment to record the responses, both correct and incorrect, in each slot separately, the counters, which recorded only total correct responses and errors per trial, were not used. Instead, the slots were numbered from one to six, and *E* did the recording by writing down the slot number for every movement of the right lever of .5 in. or more into a slot. The raw data thus permit determination of the degree of error (60°, 120°, or 180°) for every error made before entrance into the slot correct at the moment.

Subjects.—The *S*s were 78 male undergraduates from elementary psychology classes at Northwestern University. They were rewarded with one point of examination credit for each day served. None had had previous experience with the apparatus.

Procedure.—The lights and slots were paired in 25 different combinations; these were assigned to *S*s in turn until all were used, and then repeated. Sixty trials were given on the task. The first 55 trials were given on one day and the last five were given 24 hr. later. (This was done to satisfy certain requirements of a transfer task presented on the second day. The transfer data are not relevant here and will not be reported.)

Before practice began *S* was instructed to make as many correct responses per trial as

possible, and throughout practice *E* told *S* during the intertrial interval how many correct responses he had made in the previous trial.

Results and Discussion

Since we are not concerned here with learning in terms of correct responses, the acquisition curve will not be presented. Typical learning curves for this task have been presented elsewhere (1). Suffice it to say that over the 60 trials correct responses per trial increased, as usual, from a mean of about two on Trial 1 to about eleven on Trial 60.

The error data were analyzed as follows. Regardless of which slot was correct at the moment, each of the other five slots may be identified as one of the following error positions or degrees of error: 60° clockwise (CL), 120° CL, 180°, 120° counterclockwise (CCL), 60° CCL. Each time a correct response occurred, the degree of error for each of the incorrect responses preceding it was recorded. This was done for each correct response and each *S* separately; then the errors at each position were combined for all correct responses and all *S*s. The resulting data are total frequencies of errors in each error position.

It was thought that perhaps the first error made upon the appearance of each new stimulus light might be a more sensitive indicator of response generalization tendencies, so the compilation of error frequencies described above was done separately for first errors and subsequent errors. Upon plotting error gradients, such as those shown in Fig. 1, it was found that only first errors showed any significant gradient as a function of spatial similarity. Subsequent errors showed little or no evidence for a gradient; they appeared with nearly equal frequency in each of the five error positions (e.g., the frequencies of subsequent errors in each of the five positions on Trials 51–60 were: 69, 69, 57, 55, and 70). Therefore, this report will deal only with first errors. It may also be worth while noting the change that took place over practice in the percentages that first errors were of total errors (first plus subsequent errors). During Trials 1–10, 4,981 errors were made; 27.1% of these were first errors. During Trials 51–60 there were 987 errors, of which 67.6% were first errors.

The evidence for the development of response generalization tendencies, as indexed by a gradient of first errors which is directly proportional to the gradient of spatial similarity, is shown in Fig. 1. In the figure the total frequency of first errors in each of the five error positions is shown for both the first 10 and last 10 trials.

Figure 1 shows that during the first 10 trials there is no evidence for the particular shape of response

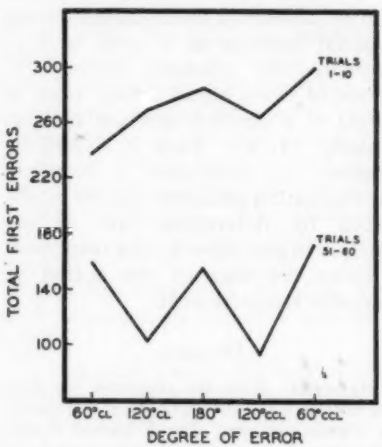

Fig. 1. Total first errors during two blocks of trials for various degrees of error both clockwise (CL) and counterclockwise (CCL) from the correct response.

generalization gradient that would be expected with this task, viz., a V or a ⌣, with the upper ends of the arms representing the 60° error positions. In fact, analysis of variance of the values at the five error positions indicates that one cannot even reject the hypothesis that the values fall along a straight horizontal line. Using the analysis of variance for correlated data, the *F* for error positions was 2.39, based on 4 and 308 *df*. This *F* value would not quite reach significance at the 5% level; it is the same value as that required at the 5% level for 4 and 400 *df*. Thus it is clear that there is no particular bias or gradient among the generalized or error responses of this task early in learning, and especially no gradient corresponding to degrees of spatial similarity.

The curve in Fig. 1 for the last 10 trials presents quite a different picture. The number of errors at each position has, of course, decreased considerably from the values for the first 10 trials. But the significant finding is that the decrease was greater for the 120° positions than for the 60° positions. Thus by a process of differential rate of extinction of errors, response generalization gradients corresponding directly to the first two degrees of similarity have been established on both sides of the correct response.

The third degree of similarity, the 180° position, is a clear-cut reversal to this trend. Concerning this, two things may be said. First, there is only one 180° slot in which all errors of the third degree of similarity must be made. Second, the diametrically opposite response may be especially subject to confusion, as in the severe confusion of right and left of some ambidextrous individuals. Thus, errors in the 180° position are greater than would be expected from the mechanism of response generalization alone. For this reason the task is not as ideal as might be for studying response generalization.

From inspection of Fig. 1 there seems no question that the curve for the last 10 trials deviates significantly from a straight line. This was confirmed by an analysis of variance of the same type as that performed earlier. The *F* was 5.80, which is significant at better than the 1% level. The *t* between the 60° CL and 120° CL positions was 2.69, between 60° CCL and 120° CCL *t* was 3.81; both values are significant at the 1% level or better. Thus, it is clear that generalized responses that were most similar to the correct response occurred as overt errors more frequently than those that were next most similar.

It is believed that the data show the development, during practice, of response generalization gradients that had not existed early in practice, even though the gradients extended over only two degrees of similarity. Further, the data indicate that the gradients developed because error responses less similar to the correct response extinguished more rapidly than those more similar. Thus, response generalization gradients may in some cases be accounted for, not in terms of parasitic reinforcement, but in terms of differential extinction rates of generalized responses that have different degrees of similarity to the correct response. This does not mean, of course, that parasitic reinforcement does not exist. The present experiment provides no test between parasitic reinforcement and differential rate of extinction.

Summary

Seventy-eight *S*s practiced a perceptual-motor task in which the responses were movements of a lever into six horizontal radial slots. Analysis of

the first error made upon the appearance of each new stimulus showed that during the first 10 trials the errors were distributed about equally in the five incorrect slots which formed a bidirectional gradient of spatial similarity around the correct slot. During the last 10 trials the frequency of errors, although considerably diminished, was directly proportional to the first two degrees of similarity, but not to the third, in both the clockwise and counterclockwise directions from the correct response. It was held that these error gradients were response generalization gradients developed by a process of differential rate of extinction of errors. A rate of extinction of errors which is inversely proportional to similarity was offered as a possible substitute for the hypothesis of parasitic reinforcement in accounting for response generalization gradients. However, it was pointed out that the experiment is not crucial with regard to parasitic reinforcement.

References

1. Duncan, C. P. Transfer in motor learning as a function of degree of first-task learning and inter-task similarity. *J. exp. Psychol.*, 1953, **45**, 1–11.
2. Hull, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
3. Morgan, R. L., & Underwood, B. J. Proactive inhibition as a function of response similarity. *J. exp. Psychol.*, 1950, **40**, 592–603.
4. Underwood, B. J. Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. *J. exp. Psychol.*, 1951, **42**, 44–53.
5. Underwood, B. J., & Hughes, R. H. Gradients of generalized verbal responses. *Amer. J. Psychol.*, 1950, **63**, 422–430.

(Received September 22, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

AVOIDANCE LEARNING TO THE ONSET AND CESSATION OF CONDITIONED STIMULUS ENERGY[1]

GEORGE BELA KISH[2]

Duke University

This study is presented as the beginning of an experimental analysis of one of the pivotal concepts in modern behavior theory, that of stimulus. In view of its basic position, it is perhaps surprising that this concept has received little experimental attention outside the field of psychophysics, and that there is no general agreement upon the characteristics of the environment necessary for functional stimulation to occur. The various definitions of the term seem to fall into two general classes. Some theories attempt a definition of physical events. Typical of such positions is that of Guthrie (4) who regards a stimulus as a "change in a physical force or a chemical state," and Hull (6) who seems to define stimulus in terms of the energy qualities of the environment present at the time of reinforcement. A second general position emphasizes the neurophysiological or behavioral consequences of stimulation. Respectively, representations of these positions can be found in the writings of Boring (2) and of Miller and Dollard (12).

Whatever the status of the second general view of stimulation, a basic problem in psychology, as Stevens (15, p. 31) has pointed out, concerns the precise specification of the physical properties of stimuli. This study attempts to evaluate the relative contributions of the two most frequently offered properties, change, and events at the time of responding. Assuming that change is the important variable in producing stimulation, one can conceive of such factors as rate, amount, and direction of change as potentially important in the specification of the attributes of a stimulus.

These experiments deal specifically with the third of these variables, direction of change, in a situation where a rat must respond to the change to avoid a shock. Assuming that the only important variable is change per se, learning should be uninfluenced by direction. If this turns out not to be the case, we are in a position to examine other possibilities including the one that some other factor is a more important element of stimulation.

METHOD

The series of four experiments involved in this report shared certain features in common. They employed the same apparatus and procedure, and each one compared the effectiveness of using the onset and cessation of stimulus energy on the rate of learning. For this reason, a general description of the experimental method is given here, and specific modifications are described in connection with the experiments to which they apply.

Apparatus and subjects.—The apparatus was a modification of that described by Kessen (9). It consisted of a wooden box with a transparent plastic front wall, an opaque plastic back wall, an electrifiable grid floor, and a wheel which protruded into *S*'s compartment and could be turned by *S*. All the opaque walls of the compartment were painted a flat white. The inside dimensions were 10 in. long, 4 in. wide, and 12 in. high. The grid bars were of $\frac{1}{8}$-in. brazing brass and were spaced at $\frac{3}{4}$-in. intervals measured from center to center of each bar.

[1] This article is based upon a Ph.D. dissertation submitted to the Graduate School of Arts and Sciences, Duke University. The author is indebted to Dr. G. A. Kimble under whose direction the research was performed.

[2] Now at the University of Maine.

The wheel was 3½ in. wide and 3¾ in. in diameter. It was made of two transparent plastic disks connected by 16 brass cross bars spaced evenly about the circumference and a ¼-in. brass rod as an axle. The wheel was mounted in a wooden housing which enclosed the outer half of the wheel and a microswitch. The microswitch was so placed that it closed each time the arm was raised by one of eight aluminum rivet heads which were inserted evenly along the circumference of one of the plastic disks. One-eighth of a turn, therefore, was adequate to close the electrical circuits controlled by the wheel.

To the Plexiglas front of *S*'s compartment was fitted a light tunnel 23½ in. long and 11 in. square. This was painted a flat white, inside, and contained a 25-w. frosted Mazda lamp centered at the far end and operated at 110 v. This arrangement provided an illumination of about 28 ft.-candles in *S*'s compartment as measured by a Weston Master II exposure meter. When the lamp was turned off, the residual illumination was not measurable either with this instrument or the Macbeth illuminometer.

The shock stimulator used in the majority of the experiments consisted of a transformer with an output voltage of 480 v. AC. This output was run through a 560,000-ohm resistor in series with *S*. The measured current delivered at the grid bars was .001 amp.

The light and shock were turned on in succession by microswitches and relays activated by a synchronous motor. The switches, relays, and the motor were housed in a sound-proof box and noises provided by an electric blower and the laboratory ventilating system served to mask other very slight noises from the apparatus.

The wheel-turning activity was recorded by a mechanical Veeder counter attached to the axle of the wheel. Latencies were measured to the nearest .05 sec. by a Standard Electric timer.

The *S*s were 72 albino rats of the Sprague-Dawley strain from 90 to 150 days of age at the time of experimentation. Upon arrival from the breeder, *S*s were housed in individual living cages where they were kept throughout the experiment on free access to food and water. They were tamed by being handled for a period of 3 to 6 days.

Procedure.—In all experiments, *S*s were run through the experimental procedure in a single session. First, *S* was placed in the experimental compartment for a 5-min. habituation period in which neither the signal nor the shock was presented. For *S*s in the "off" condition the light was present and for those in the "on" condition the light was absent. At the end of this period, Trial 1 began. Each trial consisted of the onset and continued presence of a signal (usually light or no light) which terminated with a wheel-turning response of *S*. After a specified period (5 sec. in most of the experiments), the floor of the apparatus was charged. The shock and signal continued until *S* terminated them by turning the wheel the required ¼ turn. A response occurring before the onset of shock opened the shock circuit and *S* avoided the shock on that trial. After Trial 1, *E* waited for a minimum 45-sec. intertrial interval plus a period in which no wheel turning had occurred for a period of 10 sec., and then administered Trial 2.

This cycle was repeated for 50 training trials and 30 extinction trials. The extinction trials were identical with the training trials except for the absence of shock. During extinction, if *S* did not respond to the CS, the trial was terminated manually after 30 sec. of presentation of the CS. Extinction was terminated after either five 30-sec. trials without response to the signal or after 30 trials had been presented.

Within each experiment one *S* from each group was run each day if possible and an effort was made to run *S*s in such a manner that all groups of each experiment were composed of equal numbers of *S*s which had been run at the same time of day.

The data gathered for each trial consisted of the latency of the wheel-turning response and the number of wheel turns during the 20 sec. immediately following the above response. From these data a measure of number of CR's was derived.

Experiment I

The purpose of this experiment was to test the effectiveness of the cessation of illumination as contrasted with the onset of illumination as a CS in electric shock avoidance learning. The measures of effectiveness of the two procedures were numbers of avoidance responses during training and during extinction.

Method.—The apparatus, procedure, and *S*s for this experiment did not differ from the general description given above.

Two groups of five albino rats each were randomly drawn from a pool of 60 animals. Both groups received 50 avoidance training trials with a 5-sec. interstimulus time and 30 extinction trials. For one group the CS was the onset of light and for the other group it was the cessation of light.

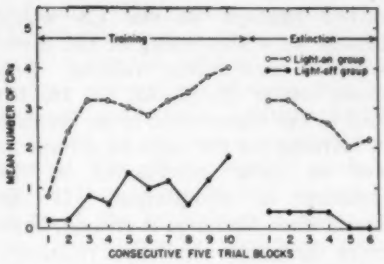

FIG. 1. Mean number of CR's in training and extinction within each block of five consecutive trials in Exp. I.

Results.—Mean number of avoidance responses (CR's with latency less than 5 sec.) is presented in Fig. 1 for each block of five training and extinction trials. It is quite clear that when the onset of light was used as the CS, conditioning occurred more rapidly than when the CS was the cessation of light.

During training, the mean total number of CR's produced by the two groups was significantly different ($t = 4.1$, $p = .005$). Because of the extreme skewness of the data for the light-off group and inhomogeneity of variance, the t test is inapplicable to the extinction measures. The fact that the six points on the extinction function are all higher for the light-on condition than for the light-off condition, however, means that, by the nonparametric sign test (**3**, p. 247), the difference is significant at the .05 level of confidence.

The wheel-turning data were, in general, extremely variable and in the same direction as the CR data, but not significant.[3]

[3] The wheel-turn data appeared to strongly reflect individual differences in general activity, responsiveness to shock, and extent of generalization between the shock and nonshock conditions. The resulting variability was so large as to make interpretation of these data for this and the remaining experiments extremely difficult. For this reason, these data are not presented.

Discussion.—These results may be interpreted as showing that direction of change is a factor of some importance influencing the effectiveness of stimulation, but they also conflict with an interpretation of the stimulus in terms of change, since if change were the factor of prime importance, a given amount of physical change should lead to similar behavioral effects regardless of the direction of that change. The following experiments attempted to shed some light upon the possible mechanism involved in producing these results.

EXPERIMENT II

The purpose of this experiment was to determine whether the effects observed in Exp. I are unique to vision. The experiment is a replication of Exp. I, comparing the onset and cessation of *auditory* stimulus energy in the same experimental situation.

Method.—The details of apparatus and procedure were the same in this experiment as in Exp. I, except that a weak buzzer was substituted for the light. For one group the onset of the noise provided by the buzzer served as the CS. For the other group the buzzer sounded continuously except for the 5-sec. interstimulus interval of silence which served as the CS. The *S*s were run in the dark.

The *S*s in this experiment were two groups of five rats each from the same pool as was used in Exp. I.

Results.—The results are presented in Fig. 2. The buzzer-on group is

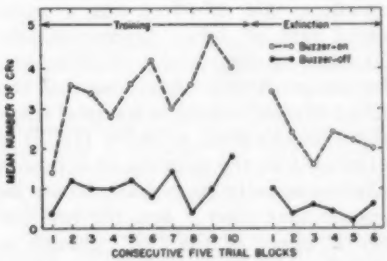

FIG. 2. Mean number of CR's in training and extinction within each block of five consecutive trials in Exp. II.

seen to be very much superior to the buzzer-off group in the number of CR's in both conditioning and extinction. The t tests comparing the mean total number of CR's in training and extinction show the buzzer-on group significantly superior. For training and extinction respectively, $t = 5.3$ and 2.83, $p = .005$ and .025.

Discussion.—These results demonstrate that the cessation of a buzzer is a less effective stimulus than the onset of a buzzer and are in harmony with the results obtained with light in Exp. I. Comparison of these data with those of Exp. I show little differences except perhaps a tendency of the buzzer-on group to be slightly superior in performance to the light-on group in training. It may tentatively be generalized that the relative ineffectiveness of the "off" type of stimulus is not a function of the sensory department stimulated.

Experiment III

A possible hypothesis to account for the results of Exp. I and II is suggested by the work of Bartley and Bishop (**1**). This work shows that the optic nerve responds not only to the onset of illumination but also to the cessation of illumination. This light-off response is termed the "off effect." A similar phenomenon is found in the auditory mechanism (**20,** p. 330). By comparison with the on-effect this off-effect has a more rapid rate of decay (approximately 50 msec. as compared with 200 msec. for the on-effect). Assuming that the speed of conditioning is maximal when the unconditioned stimulus (UCS) is presented at the moment of maximal afferent activity (**6,** p. 168), it may be argued that after 5 sec. the residual neural effect of the "off" change is considerably less than the residual neural effect of the "on" change and, therefore, that the group which received light-off as the CS would benefit by a shortening of the interstimulus time during training. Previous results (**7, 10, 13, 14, 16, 19**) lead to the expectation of an increase in learning for the light-on groups as well as those conditioned to the cessation of stimulation. If the neural after-discharge is critical, however, the effect should be relatively greater in the case of the "off" condition. This experiment was designed to provide a test of this line of reasoning.

Method.—The method did not differ from that of Exp. I except for variations in interstimulus time. Four groups of five rats were run under the conditions shown in the following diagram:

Group	*Stimulus Conditions*	*CS-UCS Interval*
I	light-off	.25 sec.
II	light-on	.25 sec.
III	light-off	1.00 sec.
IV	light-on	1.00 sec.

Together with the groups run in Exp. I, they provide three points on the function relating conditioning to interstimulus time for both the light-on and light-off conditions.

The only useful measure obtained in this experiment was the extinction measure. The interstimulus intervals were so short that few avoidance responses occurred during training. The basic data, therefore, are numbers of CR's in the 30-trial extinction period. As in the previous experiments, CR's were defined as responses with a latency of 5 sec. or less.

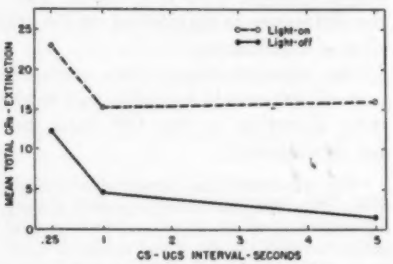

Fig. 3. Mean total number of CR's in extinction as a function of the CS-UCS interval used in training in Exp. III.

Results.—The mean total number of CR's for the various groups is graphically presented in Fig. 3 and the analysis of variance of these data is presented in Table 1. Although the variances did not prove to be homogeneous, the effects were of such a nature that the reduction in sensitivity of the analysis as a result of this inhomogeneity was not great enough to jeopardize the interpretation of the results (**11**, p. 78).

The interaction was not significant ($F = .20$). The ratio for stimulus effects was significant beyond the .005 level of confidence ($F = 17.6$). The ratio for interstimulus time effects was significant at the .05 level ($F = 3.91$).

Discussion.—The results of this experiment agree with those of the previous experiments in showing that the cessation of stimulus energy is a less effective CS than the onset of such energy. They also agree with previous results on interstimulus time in showing that the optimal interval is less than 1 sec. The lack of interaction between these two effects is not consistent with predictions based upon information concerning the off-effect of the optic nerve.

Perhaps more important than anything else in the results of this experiment is the demonstration that, under otherwise favorable conditions, learning can occur to the cessation of stimulation. The amount of learning in the "off" conditions of Exp. I and II was so small as to leave this issue in doubt. In this experiment, learning to the cessation of stimulation at the .25-sec. interval was about the same as that to stimulus onset at 1.0 sec.

TABLE 1

ANALYSIS OF VARIANCE: MEAN TOTAL CR's IN EXTINCTION

Source	*df*	Mean Square	*F*
On-off stimulus effects	1	1044	17.6**
Interstimulus time effects	2	233	3.91*
Residual	26	59.3	
Total	29		

* Significant at 5% level.
** Significant at 1% level.

EXPERIMENT IV

It has been suggested (**13**; **5**, p. 144) that the stronger the UCS in a conditioning situation, the stronger the CS required to produce maximal learning. This suggestion appears to be similar to the Yerkes-Dodson Law (**16**) which was formulated with reference to discrimination learning. This law states that the shock intensity producing optimal discrimination learning decreases as the difficulty of discrimination increases.

Assuming either that the level of illumination present at the time of reinforcement is the important element of functional stimulation or that the "off" condition is a difficult discrimination, the above principles predict that a lowering of the shock intensity would enhance conditioning to the light-off condition. Similarly, the light-on condition would be adversely affected by such a reduction in shock intensity. Statistically, results confirming this reasoning would be manifested in the form of a significant interaction between the light-on—light-off variable and shock intensity. This experiment was performed to test this prediction.

Method.—The apparatus and general procedure were the same in this experiment as in the previous ones, except that a new stimulator was substituted for that previously in use, in order to obtain a manipulable shock intensity. This shock stimulator was a Model 228 Stimulator made by C. J. Applegate and Co. of Boulder, Colorado. It made it possible to vary the shock intensity over a range extending from 0 to 5 ma. in .1 ma. steps.

Four groups of eight *S*s each were randomly selected from a general pool of 60 animals. These groups were used in a 2×2 factorial design, with shock intensities of .5 ma. and 2.00 ma. and with light-on and light-off. The

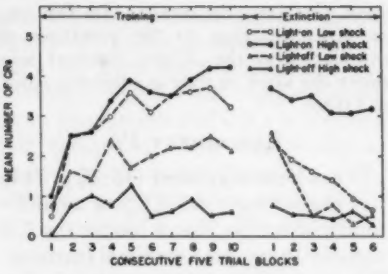

FIG. 4. Mean number of CR's in training and extinction within each block of five consecutive trials in Exp. IV.

intensity values of the lights were equal to those of the general description. The shock values were chosen on the basis of preliminary investigation.

Results.—The results in terms of mean numbers of CR's occurring within each block of five trials in training and extinction are presented in Fig. 4. The learning curves indicate that the high-shock–light-off group, as predicted, showed poorer conditioning than the low-shock–light-off group. Also as predicted, there is a tendency for the high-shock–light-on group to condition more rapidly than the low-shock–light-on group, although this tendency is not statistically significant ($t = .77$). That the effect of decreasing the shock is different for the two stimulus condition groups is shown by the results of the analysis of variance which appear in Table 2. The interaction is significant at the .025 level ($F = 5.98$). When the interaction mean square is utilized as denominator of the F ratios, neither the stimulus effects nor the shock effects are significant.

TABLE 2

ANALYSIS OF VARIANCE: MEAN TOTAL CR'S IN TRAINING

Source	*df*	Mean Square	*F*
On-off stimulus effects	1	2664	5.53
Shock intensity effects	1	144	.3
Interaction	1	481	5.98*
Within groups	28	80.46	
Total	31		

* Significant at 5% level.

The analysis of variance of the extinction measures showed no significant differences, although all values tended in the same direction as the training data. Although the statistical test of such arbitrarily selected differences cannot be justified, it is important to note that there is a large difference in extinction measures between the low- and high-shock light-on groups. The direction of this difference again supports the prediction of better learning with high shock for the light-on groups.

Discussion.—The results are consistent with the suggestion that a positive relationship exists between the CS and UCS values leading to optimal conditioning. The results are also consistent with the Yerkes-Dodson principle and suggest that these two principles are in reality special statements of a more general principle encompassing a wide variety of learning situations. The particular relevance of the results of this experiment to the problem of the nature of functional stimulation is that, to make the prediction that a lowered shock intensity would lead to an improvement in learning, it was necessary (in the case of the optimal ratio hypothesis) to assume that the CS in the situation was darkness rather than a change from light to darkness.

A possible mechanism to explain the results of this experiment may be found in the suggestion that the more intense the shock, the more behavior is elicited which is incompatible with the response to be learned. Yerkes and Dodson (16) propose that the strong shock is emotionally disorganizing. These incompatible responses would be expected to interfere more strongly with the more slowly developing, and therefore, weaker

habit which would be expected to be a function of the difficulty of the discrimination problem. This analysis must await more extended analysis of the functions of electric shock in, at least, the avoidance learning situation.

Discussion

Within the limits of the present experimental situation, the general interpretation which appears to fit the data best is that the important factor determining the effectiveness of stimulation is the amount of energy impinging upon the receptors of the organism at the time of responding or reinforcement.

In demonstrating behavioral differences between groups of animals receiving, respectively, the onset and cessation of stimulus energy, Exp. I and II indicate the inadequacy of the gross analysis of the concept, stimulus, in terms of energy change and the need for further experimental clarification. The results of Exp. IV also argue for the importance of events at the time of reinforcement. Assuming that the important aspect of the CS in the "off" condition was the low level of stimulation concurrent with the response, it was predicted that with reduced shock, learning in the light-off condition would improve. The results of the experiment confirm this prediction.

Three of the four experiments, thus, appear to minimize the importance of change in producing functional stimulation. Experiment III, however, demonstrated an appreciable amount of response strength in the "off" condition at the .25-sec. interstimulus time and implied that *S* can perhaps respond to changes in the environment over short time periods. At the .25-sec. interval the physical events at the time of the response do not differ from those at 5 sec., and yet behavioral differences as a result of these time differences are found. It would appear, therefore, that the events required for stimulation to occur require a more exact specification. Two important possibilities suggest themselves.

One of these is that the important element of stimulation is the energy at the time of the response, but that some change in the situation is necessary if *S* is to perceive the presence of these stimuli. If this interpretation is correct, it seems probable that the rate of presentation of the CS would be an important variable in conditioning. If the CS were turned on gradually, conditioning should be slower than if it is turned on abruptly.

A second likely possibility seems to be that the responses of *S* to the presence or absence of a stimulus energy are critical in determining the effectiveness of such environmental events as stimuli. If these responses are compatible with the instrumental response, it is probable that learning will be benefited. For example, in this experimental situation, turning on the light may have produced a tendency for *S* to become active, since it is known that the rat has a light-aversion drive (**8**). Turning off the light may have produced a state of relaxation incompatible with the wheel-turning response. This interpretation would handle all of the results of this study, since in Exp. III, when the CS-UCS interval was short, the shock probably appeared before the incompatible relaxation responses were complete and was presented while *S* was still responding to the light. A suggestion as to the further applicability of this line of reasoning was offered in the discussion of Exp. IV.

The present study does not provide a definitive answer to the problem of the physical nature of functional stimulation. Apparently neither the change interpretation nor the events at time of reinforcement interpretation can adequately account for the present results. It appears probable that both interpretations are partially correct. A decision with respect to this problem must await further experimentation.

Summary

Four experiments involving 72 albino rats and wheel-turning electric shock avoidance

learning were performed to explore the effectiveness of two types of stimulus: the onset of stimulus energy and the cessation of stimulus energy.

Experiment I contrasted the effectiveness of a light coming on and a light going off as stimuli. The light going off led to significantly inferior learning.

Experiment II contrasted the effectiveness of a buzzer coming on and a buzzer going off as stimuli. Similarly, significantly inferior learning was found in the buzzer-off condition.

Experiment III tested the effect of interstimulus time upon the effectiveness of the light-on and light-off types of stimulus. Beyond the usual enhancement of learning at the shorter CS-UCS time intervals, no evidence of a differential enhancement of learning with one stimulus mode over the other was found. The light-off condition still led to inferior learning.

Experiment IV tested the effect upon the two modes of stimulation of a low and high level of shock intensity. A significant interaction was shown to exist between the stimulus variables and the shock variables. Learning under the light-off condition was significantly enhanced by a *decrease* in shock intensity.

The results were discussed with reference to their bearing upon two hypotheses concerning the physical nature of functional stimulation: the hypothesis that physical energy change is the important element in stimulation and the hypothesis that the important element is the physical events at the time of reinforcement. Most of the evidence presented here appeared to favor the latter interpretation. Other possibilities, however, were discussed.

References

1. Bartley, S. H., & Bishop, G. H. Optic nerve response to retinal stimulation in the rabbit. *Proc. Soc. exp. Biol. & Med.*, 1940, **44**, 39–41.
2. Boring, E. G. *The physical dimensions of consciousness.* New York: Century, 1933.
3. Dixon, W. J., & Massey, F. J. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
4. Guthrie, E. R. *The psychology of learning.* (Rev. Ed.) New York: Harper, 1952.
5. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
6. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
7. Kappauf, W. E., & Schlosberg, H. Conditioned responses in the white rat. III. Conditioning as a function of the length of the period of delay. *J. genet. Psychol.*, 1937, **50**, 27–45.
8. Keller, F. S. Light aversion in the white rat. *Psychol. Rec.*, 1941, **4**, 235–250.
9. Kessen, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45**, 82–86.
10. Kimble, G. A. Conditioning as a function of the time between conditioned and unconditioned stimuli. *J. exp. Psychol.*, 1947, **37**, 1–15.
11. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
12. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
13. McAllister, W. R. Eyelid conditioning as a function of the CS-US interval. *J. exp. Psychol.*, 1953, **45**, 417–422.
14. Moeller, G. O., Jr. The role of the CS-UCS interval in conditioning the GSR. *Dissertation Abstr.*, 1952, **12**, 595–596.
15. Razran, G. H. S. Conditioned responses: an experimental study and theoretical analysis. *Arch. Psychol., N. Y.*, 1935, **28**, Nos. 191, 124.
16. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, **35**, 15–30.
17. Stevens, S. S., & Davis, H. *Hearing: its psychology and physiology.* New York: Wiley, 1938.
18. Stevens, S. S. (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.
19. White, C. T., & Schlosberg, H. Degree of conditioning of the GSR as a function of the period of delay. *J. exp. Psychol.*, 1952, **43**, 357–362.
20. Yerkes, R. M., & Dodson, J. D. The relation of strength of stimulus to rapidity of habit formation. *J. comp. Neurol. Psychol.*, 1908, **18**, 459–482.

(Received August 30, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

STUDIES OF DISTRIBUTED PRACTICE: XIII. INTERLIST INTERFERENCE AND THE RETENTION OF SERIAL NONSENSE LISTS

BENTON J. UNDERWOOD AND JACK RICHARDSON

Northwestern University[1]

Hovland (1) found retention to be better following distributed practice than following massed practice of serial nonsense lists. With roughly comparable materials, we have consistently found massed practice to produce superior retention (**6, 7, 10**). One factor which appeared to be at least partially responsible for this contradiction is the ability level of *Ss*. We have shown (**11**) that slow-learning *Ss* give better retention following massed practice than following distributed, while little influence of intertrial interval is apparent in the retention scores of fast-learning *Ss*. Since Hovland's *Ss* as a group were more rapid learners than even our fast-learning group, it is quite possible that distributed practice facilitates retention only for very fast-learning *Ss*. In the present investigation, however, we wish to analyze another consequence of the differences in the practice level of Hovland's and our *Ss*.

In our studies on retention of paired-associate lists we have usually found that retention following distributed practice is better than retention following massed (**5, 8, 9**). While the differences were not always significant statistically, they were almost always in the same direction. Now, we have shown elsewhere (**13**) that interlist interference effects are considerably greater for paired-associate learning than for serial learning. A consideration of these facts, namely, that retention is more likely to be facilitated by distributed practice of paired-associate lists than for serial lists, and that interlist interference effects are greater for paired-associate than for serial lists, suggests another reason for the contradiction between Hovland's findings and ours. In Hovland's studies *Ss* served for 22 days and learned 22 different lists of syllables. Because of the very limited number of letters available for constructing syllables, the interlist interference built up in Hovland's studies would be much greater than in ours, where *S* learned no more than four lists. In view of these facts, our prediction is that distributed practice will facilitate retention of serial lists only when interlist interference is high, and that massed practice will facilitate retention when interlist interference is low. These predictions are derived entirely from consideration of available facts.

In the present experiments we have constructed short consonant lists having high interlist similarity. Our expectation is that for the first-learned list, massed practice will produce better retention than distributed practice, but that after learning a series of lists so that interlist interference is high, distributed practice will produce better retention.

[1] This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and The Office of Naval Research.

METHOD

General procedure.—Each *S* learned seven serial lists of consonant syllables. In learning List 1, half of the *S*s were given massed practice (2-sec. intertrial interval) and half distributed practice (30-sec. interval). Retention of List 1 was measured after 24 hr. On this second day, following the relearning of List 1, all *S*s learned two more lists by massed practice. On Day 3 all *S*s learned three additional lists by massed practice. On Day 4, half the *S*s learned List 7 by massed practice and half by distributed practice. The retention of List 7 was taken on Day 5, 24 hr. after learning. According to our expectations, retention of List 1 should be better following massing than following distribution, and retention of List 7 should be better following distribution. Lists 2 through 6 were inserted in the schedule entirely for the purpose of building up heavy interlist interference prior to learning the last (seventh) list.

Lists.—Each of the seven serial lists used was made up of six consonant syllables taken from Melton (**3**). All syllables were between 50% and 75% association value as calibrated by Witmer (**14**). Since the first syllable was used as an anticipatory cue, only five items were actually learned. A total of nine consonants appear in the syllables, namely, C, H, F, K, M, Q, R, T, and X. Each consonant appeared twice in each list for a total of 14 times all lists considered. The seven lists are reproduced below.

LISTS

A	*B*	*C*	*D*	*E*	*F*	*G*
CRX	HQM	XRK	CQM	FXR	RKX	MRH
HTQ	MCK	CHM	KRX	HMQ	QTH	FQT
RFM	HFR	QMT	XTH	CHX	MXF	RKC
QCK	XCT	FRX	HRF	MTQ	CMT	XHT
MKX	FRK	KFT	QMC	KCT	HQF	FQX
HFT	TXQ	HQC	FTK	RKF	CKR	CMK

It may be worth noting that lists composed of these syllables are incredibly difficult to learn. In pilot work we first used 11 items, then 7, and finally the 6-item lists presented here. Even so, we lost a few *S*s who could not learn one list within a 50-min. period.

Lists were presented on a Hull-type drum at a 2-sec. rate, with the anticipation method used at all times. The learning of each list was carried to one perfect recitation.

Design details.—A total of 168 undergraduates served as *S*s. Half of these ($N = 84$) always served under massed conditions, that is, learning of all seven lists took place with a 2-sec. intertrial interval. The other half all had a 30-sec. intertrial interval for List 1, massed practice for Lists 2 through 6, and distributed practice for List 7. For List 7, 28 *S*s learned with a 1-min. distribution interval, 28 with a 2-min. interval, and 28 with a 3-min. interval. Thus, we have four groups: a massed group of 84 *S*s, and three groups of 28 *S*s each having varying degrees of distributed practice on List 7. For all *S*s the relearning of Lists 1 and 7 was by massed practice 24 hr. after original learning.

The positions of the seven lists in the series were systematically varied so that each occurred at each position four times in a block of 28 *S*s. Standard instructions for anticipation learning were given prior to learning the first list, and *S*s were always urged to recall as many words as possible at the time of the relearning of the first and seventh list. Intertrial intervals during distributed practice were filled with symbol cancellation (**6**).

RESULTS

Acquisition.—The mean number of trials required to learn each list is plotted in Fig. 1. On List 1 we have only two basic groups, the massed group and the 30-sec. group, each of 84 *S*s. On Lists 2–6 all *S*s were given massed practice. On List 7 there is the massed group and the three distributed groups (1-min., 2-min., and 3-min. intertrial intervals.)

We may note initially that on List 1 distributed practice facilitated markedly. On Lists 2–6 the two basic groups are roughly comparable, as they should be since conditions were the same for both groups and since deviant *S*s were replaced so that

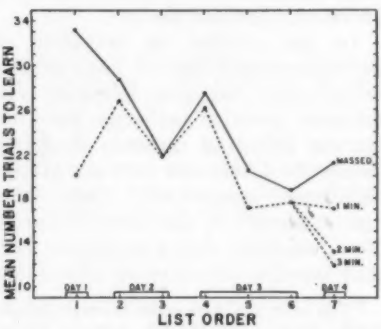

FIG. 1. Mean number of trials to attain one perfect recitation on each of the seven lists.

the massed group and the three distributed groups would be comparable on these lists. Between-day variations are apparent in learning, especially between Days 2 and 3. List 3 was the last list learned on Day 2 and List 4 the first on Day 3. As can be seen, List 4 was learned at a considerably slower rate than List 3. This probably can be attributed to at least two factors, namely, lack of warm-up and interference from previous lists. This latter possibility is suggested by the fact that there was a clear-cut increase in overt errors per trial on List 4 as compared with List 3.

In learning List 7 all three distributed groups learned more rapidly than did the massed group. Heterogeneity of variance prevents simple statistical tests, and since the finding that distributed practice facilitates the acquisition of serial nonsense lists merely confirms many previous findings, we have not transformed the scores. If we ignore the heterogeneity, the *F*'s are highly significant for the scores on both the first and seventh list.

It is important, however, to determine whether or not the three distributed groups differ in rate of learning the seventh list. Figure 1 suggests that speed of learning and length of intertrial interval are directly related. In previous studies (**6, 7, 10, 15**), under conditions of relatively low interlist interference, length of interval beyond a certain point (approximately 1 min.) has not been a relevant variable. The scores for the three distributed groups on List 7 were log transformed to eliminate heterogeneity of variance. The *F* for these transformed scores was 4.12. With 2 and 80 *df*, *F* at the 5% level is 3.11 and at the 1% level, 4.88. In view of the direct relation between

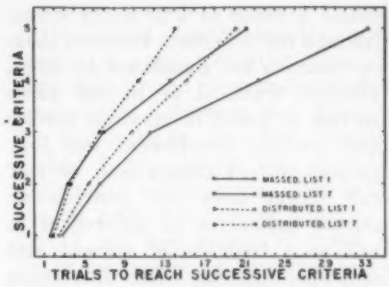

FIG. 2. Acquisition curves for Lists 1 and 7 under massed and distributed practice.

length of interval and speed of learning as seen in Fig. 1, and in view of the retention results to be presented shortly, we conclude that when interlist interference is high, facilitation in learning by distributed practice is a direct function of the length of the distribution interval.

One final set of data will be presented concerning acquisition. We have seen that distributed practice facilitated the acquisition of both Lists 1 and 7. However, there is some indication that processes responsible for this facilitation are not necessarily the same for Lists 1 and 7. Figure 2 shows trials-to-criteria curves for the massed and distributed groups on List 1 and List 7. For List 7 the three distributed groups have been combined. It will be noted that in learning List 1 the distribution of practice produces superiority throughout the entire course of learning. But, in learning List 7, distribution results in initially somewhat inferior performance followed by a rapid change to superior performance.

Recall of List 1.—Retention of List 1 was measured 24 hr. after learning. The 84 *Ss* who learned with massed practice recalled a mean of 2.35 ± .15 items, while those 84 *Ss* who learned with a 30-sec. intertrial interval

recalled a mean of 2.03 ± .14 items. Although the difference between these two values is not significant by usual statistical tests, it is in the same direction as found in previous studies when interlist interference was low. However, direct comparison of raw recall scores does not provide an accurate indication of differences in recall as a function of massed and distributed learning when differences in rate of acquisition are as marked as they are here.

We have shown elsewhere (12) that when materials are learned to the same criterion but at different rates, direct comparison of recall scores is not justified. We also presented a solution to this problem, a solution which we termed a successive probability analysis. When this technique was applied to the present learning and recall data of List 1, it was quite apparent that the massed items were better recalled than the distributed items. In the paper in which we explained in detail the method for making a successive probability analysis, we did not consider a technique for making statistical tests of significance for the resulting recall data. Such a technique has now been worked out by one of the present authors (JR). While we cannot describe this method in detail in the present paper, the general approach can be given. Essentially, an expected recall score is obtained for each *S* if recall had been taken immediately after learning. This expected score is obtained by using the group curve which describes the successive probability function. Then, the obtained raw recall score after 24 hr. for each *S* is noted, and the difference between the expected and obtained score taken. Thus, for each *S* we have a difference score. When conditions are to be compared, an analysis of variance is performed on these distributions of difference scores. Such an analysis has been made of the difference scores for recall of List 1 following massed and distributed practice and a highly significant *F* was found, with retention better following massed practice than following distributed practice. In view of the above considerations, and in view of previous findings in this series (6, 7, 10), we feel justified in concluding that when interlist interference is low, massed practice results in better retention than does distributed practice.

Recall of List 7.—The mean number of List 7 items recalled by each group is shown in Fig. 3. Recall is directly related to length of intertrial interval. Of the 84 *S*s who learned under the massed condition, 64% were unable to give any correct items on the recall trial. Of the 28 *S*s who learned with a 3-min. intertrial interval, only 14% were unable to give at least one correct response at recall.

We discussed earlier the problems attendant to an evaluation of differences in recall when material is learned at different rates to the same criterion. A comparison of learning of List 7 (Fig. 1) and its recall (Fig. 3) shows that rate of learning and recall are directly related. Are the differences in Fig. 3 merely a reflection of the differences in rate of learning or do they represent true differences in response strength? The evidence indicates that there are true differences in response strength. Our method for arriving at this conclusion is the difference-score method as discussed earlier. We derived two distributions, one for the 84 *S*s having massed practice and one for the 84 having distributed practice. The score for each *S* was the difference between the expected and obtained score, the expected score being based on the group probability function. A critical ratio between the means of these two

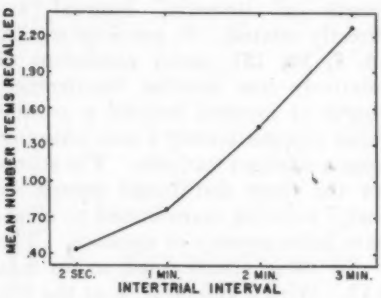

FIG. 3. Recall of List 7 as a function of intertrial interval.

distributions gives a value of 4.67, which is highly significant. The difference scores for the massed group were considerably skewed so that normal probability statistics may not be entirely appropriate. However, use of the median test as suggested by Moses (4) gives a chi-square value of 9.52, which is significant well beyond the .01 level for 1 *df*.

In summary of the recall data, the evidence indicates that massed practice is followed by better recall than is distributed practice when interlist interference is low (List 1), and just the reverse is true when interlist interference is high (List 7).

Relearning.—In general, the relearning scores reflect the differences found in recall as presented above. For List 1, a mean of 5.50 ± .46 trials was required to relearn following massing, while a mean of 7.06 ± .56 trials was needed following distribution. The *F* is 4.26, with a value of 3.91 needed for the .05 level of confidence. In relearning List 7 the mean values were 9.10 ± .52, 7.29 ± .82, 6.04 ± .59, and 5.50 ± .74, for the 2-sec., 1-min., 2-min., and 3-min. intertrial intervals, respectively. The *F* for these means (6.85) is significant beyond the .01 level of confidence. Considering only the three distributed conditions *F* is 1.62, which falls considerably short of the .05 level. Thus, while the relearning means show an inverse relationship with intertrial interval, the length of the interval beyond 1 min. does not produce significant differences in relearning. Nevertheless, all of these data present essentially the same picture as did the recall scores, namely, with low interlist interference massed practice results in better retention than does distributed practice while with high interlist interference the reverse is true.

Fast vs. slow Ss.—A previous paper (11) showed that massed practice results in better retention than distributed practice for slow-learning *Ss*, while little difference was apparent for fast-learning *Ss*. For retention of List 1, the present data show this same set of relationships. We divided the 84 *Ss* having massed practice on List 1 into two groups of equal *N*'s based on speed of learning List 1. The same was done for the 84 *Ss* having distributed practice. For slow *Ss*, following massed practice of List 1, a mean of 2.76 items was recalled and 5.24 trials were required to relearn. Following distributed practice slow *Ss* recalled 1.93 items and required 8.88 trials to relearn. Fast *Ss*, on the other hand, showed slightly better retention following distributed practice than following massed. Massed learning resulted in a mean recall of 1.95 items and 5.76 relearning trials. For distributed practice, 2.14 items were recalled and 5.24 trials were required to relearn. Thus, these data are essentially the same as those presented in the earlier paper. Slow *Ss* give better retention following massed practice while fast *Ss* show little difference in retention following massed or distributed learning.

However, when we turn to retention of List 7, where interlist interference is high, the differences in retention following massing and distribution as a function of ability level disappear entirely. Recall and relearning curves for fast and slow *Ss* as a function of intertrial interval show no differences of any consequence. For the analysis of retention of both List 1 and List 7 as a function of ability level and intertrial interval a successive probability analysis is indicated because of differences in slopes of acquisition curves. However, the lack of sufficient number of cases for such a

breakdown made this impossible. Nevertheless, adjustments which would be made by such an analysis would not change the relationships presented here for fast and slow *Ss*, although the magnitude of the differences may change. We conclude, therefore, that while ability level interacts with intertrial interval when interlist interference is low, it does not do so when interlist interference is high.

Discussion

The results have shown that massed practice of serial lists results in better retention than does distributed practice when interlist interference is low, but that when interlist interference is high, distributed practice results in better retention. It is our conclusion that Hovland's (1) finding that retention was better following distributed than following massed practice resulted from the high interlist interference built up by the extended schedule of learning undertaken by his *Ss*. In previous studies we have tended to find better retention of paired associates following distributed practice than following massing. We have also shown that paired-associate lists interact much more than do serial lists (13). It is our conviction, therefore, that distributed practice facilitates retention only when a certain amount of interlist interference is present.

We shall not attempt in this paper to present any complete theory to incorporate the known facts of massing and distribution. However, we do wish to discuss briefly certain notions which we have found useful in planning experiments undertaken since the data of the present experiment were first analyzed.

When interlist interference is low, the length of intertrial interval beyond a certain minimum (possible 60 sec.) is not an important variable in either acquisition or retention. While this fact is fairly clear from previous work (as noted earlier), we ran a special group of 14 *Ss* under the conditions of the present experiment but giving a 2-min. intertrial interval on the first list. The retention scores for this group were no higher than for the group having the 30-sec. interval. We have also noted that when interlist interference is low, slow *Ss* tend to recall better following massed practice than following distributed practice, whereas fast *Ss* are likely to show little difference. We believe, therefore, that some type of inhibition theory may adequately handle the facts when interlist interference is low. This inhibition would need to develop during learning and dissipate very rapidly with rest. Under massed practice the inhibition has little opportunity to dissipate, hence performance is depressed. Over a retention interval, however, the inhibition dissipates leaving the massed list apparently stronger than the list learned by distribution to the same criterion. Obviously, in one form or another, such an inhibition theory as suggested in its general form here has been proposed by many previous writers to account for certain phenomena both in motor and verbal learning as well as in conditioning. In any event, it does seem to handle adequately the general facts of massing and distribution when interlist interference is low.

However, when we turn to the data obtained when interlist interference is high, a simple inhibition theory seems hopelessly inadequate. Our present notion is that inhibition as discussed above either develops less and less as practice on new lists continues or is so overshadowed by additional processes as a consequence of the high interlist interference that it becomes a minor matter. That the inhibition might actually develop in smaller and smaller amounts as practice continues would have some support from motor-learning data (2). But, whatever the case may be, it seems apparent that an additional construct or constructs must be admitted to handle the situation where interlist interference is high. We may note two facts which we think suggest the general nature of the characteristics which must be assigned the construct. First, length of

intertrial interval becomes a much more significant variable than when interlist interference is low. Secondly, the acquisition curves indicate that some process is involved which initially delays learning by distributed practice when interlist interference is high. This is not true when interlist interference is low. It may be mentioned that in data (as yet unpublished) on learning of paired-associate lists this early retardation in learning by distributed practice when interlist interference is high is more pronounced than in the present acquisition curves for serial lists. These facts suggest to us that during the early rest intervals in distributed practice there is a loss of differentiation among response systems so that when a learning trial is given following rest, interference retards learning (as compared with massing). However, the re-establishment of differentiation following its loss, will, with continued alternating rest and work periods, allow for a more permanent "extinction" of the interference effects. We would also have to say that loss of differentiation is directly related to length of intertrial rest, but that the greater the loss (the greater the recovery of interference) the more permanent the extinction of the interfering tendencies. This would not only allow for ultimate acquisition rate to be directly related to length of intertrial interval, but would also allow retention to be directly related to this interval.

These crude notions allow certain gross predictions concerning the effect of length of retention interval, amount of interlist interference, length of intertrial interval, as well as predictions concerning retroactive inhibition with distributed practice on interpolated learning. We hope to test some of these predictions and if the results prove encouraging a sharpening or refinement of the theory will obviously be necessary.

Summary

A consideration of available facts made it seem likely that massed practice should facilitate retention when interlist interference is low and distributed practice when interlist interference is high.

The materials used to test these predictions were seven serial lists of consonant syllables with high interlist similarity. The experimental schedule for each *S* extended over five days. Half of the 168 *S*s were given massed practice on List 1 and half distributed practice. Five additional lists were then learned by all *S*s under massed practice. On List 7 half were given massed practice and half distributed, the latter employing three subgroups having 1-, 2-, and 3-min. distribution intervals. Retention of Lists 1 and 7 was measured after 24 hr.

The results show that distributed practice facilitated acquisition of both Lists 1 and 7, and this facilitation was related directly to length of intertrial interval on List 7. Retention of List 1 was better following massed practice than following distributed practice, while the reverse was true for List 7. Facilitation in retention of List 1 by massed practice was largely confined to slow-learning *S*s, but ability level was not a variable in retention of List 7.

It was concluded that Hovland's findings that distributed practice facilitates retention was due to the heavy interlist interference which existed as a consequence of the large number of syllable lists learned by his *S*s. Our previous studies, as well as the present data, indicate that massing produces superior retention when interlist interference is low.

The data indicate that at least two theoretical notions will be required to handle the facts. When interlist interference is low, a simple inhibition theory may be adequate. When interlist interference is high, an additional construct will be needed. The characteristics required of this construct were indicated.

References

1. Hovland, C. I. Experimental studies in rote-learning theory: VI. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, **26**, 568–587.
2. Kimble, G. A., & Shatel, R. B. The relationship between two kinds of inhibition and the amount of practice. *J. exp. Psychol.*, 1952, **44**, 355–359.
3. Melton, A. W. Materials for use in experimental studies of learning and retention of verbal habits. Mimeo. manuscript, Univ. of Missouri, 1940.
4. Moses, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122–143.

5. Underwood, B. J. Studies of distributed practice: II. Learning and retention of paired-adjective lists with two levels of intra-list similarity. *J. exp. Psychol.*, 1951, **42**, 153–161.

6. Underwood, B. J. Studies of distributed practice: VI. The influence of rest-interval activity in serial learning. *J. exp. Psychol.*, 1952, **43**, 329-340.

7. Underwood, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intra-list similarity. *J. exp. Psychol.*, 1952, **44**, 80–87.

8. Underwood, B. J. Studies of distributed practice: VIII. Learning and retention of paired nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 133–142.

9. Underwood, B. J. Studies of distributed practice: IX. Learning and retention of paired adjective lists as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 143–149.

10. Underwood, B. J. Studies of distributed practice: XI. An attempt to resolve conflicting facts on retention of serial nonsense lists. *J. exp. Psychol.*, 1953, **45**, 355–359.

11. Underwood, B. J. Studies of distributed practice: XII. Retention following varying degrees of original learning. *J. exp. Psychol.*, 1954, **47**, 294–300.

12. Underwood, B. J. Speed of learning and amount retained: A consideration of methodology. *Psychol. Bull.*, 1954, **51**, 276–282.

13. Underwood, B. J. Intralist similarity in verbal learning and retention. *Psychol., Rev.*, 1954, **61**, 160–166.

14. Witmer, L. R. The association value of three-place consonant syllables. *J. genet. Psychol.*, 1935, **47**, 337–360.

15. Wright, S. T., & Taylor, D. W. Distributed practice in verbal learning and the maturation hypothesis. *J. exp. Psychol.*, 1949, **39**, 527–531.

(Received September 20, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

THE ROLE OF IRRELEVANT STIMULI IN HUMAN DISCRIMINATION LEARNING[1]

MORTON HAMMER

State University of Iowa

In discrimination problems the organism is generally confronted with a variety of stimuli of which some are relevant and some are irrelevant. Relevant stimuli are those whose presence is correlated with the reinforcement and nonreinforcement of the discriminative responses, and to which the appropriate discriminative responses must become connected if the discrimination is to be learned. No such relationships obtain for the irrelevant stimuli.

Discrimination-learning theorists have recognized that, during discrimination training, the organism learns responses that increase the likelihood of his being stimulated by the relevant stimuli. This learned behavior has been referred to as receptor orienting behavior (**5, 6**), observing responses (**7**), and as a mediating process affecting the distinctiveness of cues (**3, 4**). For convenience the indifferent term "attend to" will be used to encompass the kind of behavior implied by these concepts.

It has been suggested further that during discrimination training there is a reduction in the frequency with which the organism attends to the irrelevant stimuli (**1, 4**). This reduction may be due to one of the following alternatives: (*a*) As the organism learns to attend to the relevant stimuli it attends less to the irrelevant stimuli simply because attending to one conflicts with attending to the other. (*b*) The organism learns *not* to attend to the irrelevant stimuli, and, following discrimination training, will be less likely to attend to the irrelevant stimuli regardless of the presence of the relevant stimuli.

Where a decrease in attending to the irrelevant stimuli has been referred to in the literature, it generally has not been made explicit which of these alternatives is being suggested. Lawrence (**4**), for example, stated that his *S*s "learned to 'ignore' " certain irrelevant cues, but he did not indicate the nature of this learning. Krechevsky (**2**), however, has stated that the organism selects out of the situation certain stimuli to which he attends, and that he continues to attend to these stimuli until he learns that they are irrelevant, whereupon he gives up responding to these stimuli and selects another set of stimuli to respond to. Krechevsky appears to be asserting that the organism does learn not to attend to the irrelevant stimuli.

The purpose of the present experiment was to determine whether *S* learns not to attend to the irrelevant stimuli in discrimination situations. A transfer of training design was used. The *S* was presented successively with two discrimination problems. The problems were similar in kind, each requiring *S* to discriminate between

[1] This article is based on a dissertation submitted to the faculty of the Department of Psychology of Indiana University in partial fulfillment of the requirements for the Ph.D. degree, September, 1953. The author wishes to express his appreciation to Drs. H. G. Yamaguchi, chairman of the thesis committee, and W. K. Estes for their assistance in conducting this research; and to Drs. I. E. Farber and L. D. Goodstein for their helpful criticism during the preparation of this article.

two classes of stimulus situations containing both relevant and irrelevant stimulus elements. Stimulus elements that were irrelevant in the first problem were used in the second problem as irrelevant elements for some *S*s, and as relevant elements for others. The second problem also included stimulus elements that had not appeared in the first problem.

It was expected that if *S* had learned, in the first problem, not to attend to the irrelevant elements, and if this learning did not depend upon the presence of the relevant elements, his performance on the second problem would be affected in the following way: (*a*) When the irrelevant elements were retained as irrelevant elements, *S*s would attend more readily to the new relevant elements and learning would be facilitated. (*b*) When irrelevant elements were used as relevant elements in the second problem, *S* would be less likely to attend to them and learning would be retarded.

Method

Apparatus.—A discrimination panel with one stimulus exposure window and two response keys was used. The keys were at the base of the panel directly in front of *S*. On the panel below the window was a red light bulb that could be connected in series with either of the keys. A shutter located behind the window was operated manually by *E*. When *E* raised the shutter the stimulus material was exposed to *S*.

Stimulus materials.—The stimulus elements were capitalized consonants. Eighteen consonants were divided randomly into the following six "subclasses" of elements: 1, YQD; 2, FPN; 3, WTJ; 4, LRS; 5, KXM; 6, VZC. A stimulus class was composed of two subclasses and designated by the numbers of those subclasses indicated above. For example, Stimulus Class 1–5 was composed of consonants Y Q D K X M. On any one trial only four of the six consonants were presented, two consonants from each subclass. The four consonants were typed on 3 × 5-in. white index cards in a diamond pattern. A stimulus presentation from Stimulus Class 1–5 could be the following, with D and Q from Subclass 1, and K and M from Subclass 5:

Q
D K
M

Experimental design.—Each *S* was presented with two discrimination problems. In their first problem, *S*s of Experimental Group Ia were required to discriminate between Stimulus Class 3–1 and Stimulus Class 2–1. Subclass 1 was shared by both stimulus classes and so provided irrelevant elements, whereas Subclasses 2 and 3 were each unique to a single stimulus class and so provided relevant elements. On each trial a stimulus card was presented containing four consonants, two from Subclass 1 and two from either Subclass 2 or Subclass 3. The *S* responded by pressing one of the two keys. If the consonants were from Stimulus Class 3–1, a left key press was correct; if the consonants were from Stimulus Class 2–1, a right key press was correct.

After learning the first problem, *S* was presented with a second problem similar to the first but involving two new stimulus classes. In their second problem *S*s of Experimental Group Ia were required to discriminate between Stimulus Classes 1–5 and 4–5. Subclass 1 appeared in both problems for this group of *S*s. In the first problem it provided irrelevant elements and in the second problem relevant elements.

The only difference in treatment between experimental and control groups was a difference in the irrelevant elements of the first problem, as is evident in Table 1. For their first problem *S*s

TABLE 1

Outline of Experimental Design Indicating by Their Numerical Designations the Stimulus Classes Involved in Each Problem

Group		1st Problem		2nd Problem	
		L*	R	L	R
Ia	Experimental	3–1	2–1	1–5	4–5
	Control	3–6	2–6	1–5	4–5
Ib	Experimental	3–6	2–6	6–5	4–5
	Control	3–1	2–1	6–5	4–5
IIa	Experimental	3–5	2–5	1–5	4–5
	Control	3–6	2–6	1–5	4–5
IIb	Experimental	3–6	2–6	1–6	4–6
	Control	3–5	2–5	1–6	4–6

* "L" and "R" indicate which key (left or right) was to be associated with a given stimulus class.

of Control Group Ia were required to discriminate between Stimulus Classes 3–6 and 2–6. Their second problem was the same as that of Experimental Group Ia, however, for the control *Ss* all of the stimulus elements in the second problem were new. It was considered possible that the difference in irrelevant elements of the first problem might make the control *Ss'* first problem more or less difficult than that of the experimental *Ss*, and that this difference in difficulty might alone account for differences in the performance of experimental and control *Ss* on the second problem. To partially control for this possibility Experimental and Control Groups Ib were introduced and a counterbalancing procedure was used which provided that a given discrimination task would be presented to half of the experimental *Ss* and half of the control *Ss*. The first problem for Experimental Group Ib was identical with that of Control Group Ia; and the first problem of Control Group Ib was identical with that of Experimental Group Ia, as Table 1 shows.

Similar discrimination problems and controls were used for Groups IIa and IIb. The important differences among the experimental and control groups are the following. For experimental Groups Ia and Ib the irrelevant elements of the first problem were used in the second problem as a subclass of relevant elements to be associated with a left key press. The other subclass of relevant elements and the irrelevant elements for the second problem were new. For Experimental Groups IIa and IIb the irrelevant elements of the first problem were retained as irrelevant elements in the second problem. The relevant elements were new. For Control Groups Ia, Ib, IIa, and IIb none of the stimulus elements used in the first problem appeared in the second problem. Since Control Groups Ia and IIa were required to discriminate between the same stimulus classes, a single group of *Ss* represented both. Ten *Ss* were included in each group, making a total of 40 experimental *Ss* and 30 control *Ss*. All were obtained from undergraduate psychology courses at Indiana University.

Experimental procedure.—A pretraining stage was included to facilitate discrimination learning. The *Ss* were shown 18 cards from each of the two stimulus classes they were to encounter in the first discrimination problem. They were told that the cards from one stimulus class belonged to Set I and the cards from the other stimulus class belonged to Set II, and were instructed to try to distinguish between the two sets of cards. The *E* presented the cards one by one at the stimulus exposure window, in alternating blocks of nine cards each from the two stimulus classes. The *Ss* were told whether the cards belonged to Set I or to Set II. These pretraining cards were not used in the remainder of the experiment. The exposure time for each card was approximately 5 sec.; intertrial interval was approximately 4 sec.

After pretraining, *Ss* were instructed to press one key when they thought a card belonged to Set I and to press the other key when they thought a card belonged to Set II. They were informed that the red light would flash on if they pressed the correct key and that they would continue to work at the problem until they made a certain number of consecutive correct responses. The exposure time varied between 2 and 3 sec. Intertrial interval was approximately 6 sec. The stimulus cards were reshuffled for each *S*. The order of presentation of the two stimulus classes differed for each *S* and was determined randomly, except that within a block of 10 trials each stimulus class was presented five times. The first problem was continued until *S* made 20 consecutive correct responses. The maximum number of trials possible was 144 since after pretraining 72 cards were available for each stimulus class. The *Ss* not meeting the criterion of 20 consecutive correct responses were rejected. As soon as the criterion was me., the second problem was begun without any interruption or change in procedure. The *Ss* received 50 trials on the second problem.

Results and Discussion

Each *S* received an error score which was the number of incorrect key presses made during the 50 trials of the second problem. Since *t* tests failed to indicate significant differences between lettered subgroups, the letter designations were dropped, leaving Experimental Groups I and II, and Control Groups I and II.

It was expected that if *Ss* learned not to attend to the irrelevant elements, the error scores of Experimental Group I would exceed those of Control Group I, and the error scores of Experimental Group II would be less than those of Control Group II. The means, variances, and *t*-test results are presented in Table 2. The results indicate that the groups did not differ significantly. The mean error scores for both experimental groups exceeded those of the

TABLE 2

DIFFERENCES BETWEEN EXPERIMENTAL AND CONTROL GROUPS IN MEAN ERROR SCORES ON THE SECOND PROBLEM

Group	*N*	Means	Variance	*t*	*P*
Exp. I	20	13.60	48.88	1.17	>.20
Con. I	20	11.35	25.71		
Exp. II	20	13.65	42.66	1.28	>.20
Con. II	20	11.10	36.73		

control groups so that the over-all differences between groups were not in the expected direction.

Since the transfer of irrelevant stimulus elements from the first problem to the second problem did not significantly affect *S*'s performance on the second problem, there was no evidence that responses of any sort were acquired to the irrelevant elements during discrimination training. It is important to note that there were irrelevant stimuli present besides those experimentally defined as irrelevant, e.g., apparatus and proprioceptive stimuli. Very likely these stimuli did serve as eliciting stimuli for responses that oriented *S* towards the stimulus exposure window. If the experimentally defined irrelevant elements could have been distinguished from the relevant elements on the basis of location, it is possible that similar orienting responses would have become attached to some of them, and these S–R connections might have produced differences between the experimental and control groups in the second problem. The present study suggests that where location is not a distinguishing stimulus variable, responses are not associated with the irrelevant elements during discrimination training.

SUMMARY

The purpose of the present experiment was to determine whether *S* learns *not* to attend to the irrelevant stimuli during discrimination training. The *S* was presented successively with two discrimination problems. A transfer of training design was used. For Experimental Group I irrelevant stimuli from the first problem were used as relevant stimuli in the second problem. It was expected that if *S* had learned, in the first problem, not to attend to the irrelevant stimuli, *S* would be less likely to attend to them in the second problem, where they were relevant, and learning would be retarded. For Experimental Group II, irrelevant stimuli from the first problem were used as irrelevant stimuli in the second problem. It was expected that if *S* had already learned not to attend to the irrelevant stimuli, he would more readily attend to the relevant stimuli in the second problem and learning would be facilitated.

The expected results were not obtained. The transfer of irrelevant stimuli from the first problem to the second problem did not significantly affect *S*'s performance on the second problem and there was no evidence that responses of any sort were acquired to the irrelevant stimuli during discrimination training.

REFERENCES

1. BUSH, R. R., & MOSTELLER, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.
2. KRECHEVSKY, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, **45**, 107–133.
3. LAWRENCE, D. H. Acquired distinctiveness of cues: I. Transfer between discrimination on the basis of familiarity with the stimulus. *J. exp. Psychol.*, 1949, **39**, 770–784.
4. LAWRENCE, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *J. exp. Psychol.*, 1950, **40**, 175–188.
5. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
6. SPENCE, K. W. Continuous versus noncontinuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, **47**, 271–288.
7. WYCKOFF, L. B. The role of observing responses in discrimination learning. Part I. *Psychol. Rev.*, 1952, **59**, 431–442.

(Received October 4, 1955)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

TWO TESTS OF THE SHEFFIELD HYPOTHESIS CONCERNING RESISTANCE TO EXTINCTION, PARTIAL REINFORCEMENT, AND DISTRIBUTION OF PRACTICE [1]

WILMA WILSON, ELIZABETH J. WEISS,[2] AND ABRAM AMSEL

Tulane University

Sheffield (**7**) has offered a stimulus generalization interpretation of the finding that responses acquired under partial (50%) reinforcement are more resistant to extinction than those acquired under continuous (100%) reinforcement conditions. She assumes that some part of the stimulation, on any trial after the first in a series of trials, is composed of carried-over traces of goal stimulation from the previous trial; so that stimulation on trials following reinforcement would be unlike that during any extinction trial after the first, while stimulation on trials following nonreinforcements, in partial reinforcement training, would be similar to that during extinction trials. Consequently, there should be more generalization decrement from 100% reinforcement to extinction than from 50% reinforcement to extinction, and resistance to extinction following partial reinforcement should be greater than after continuous reinforcement.

Sheffield reasoned that if time were allowed between acquisition trials for carry-over stimulation to dissipate, increased resistance to extinction following partial reinforcement should be eliminated. In an experimental test of this interpretation, Sheffield found that 50% reinforcement resulted in greater resistance to extinction than 100% reinforcement after massed but not after spaced acquisition.

Several variations of the Sheffield experiment have been performed using different response systems or different time intervals (**3, 4, 11**). In each case, the partial reinforcement groups were more resistant to extinction than the continuous ones. Sheffield's finding, that partial and continuous groups were equally resistant to extinction following spaced training, has not been confirmed. Her results have been considered the most significant exception to the usual finding of increased resistance to extinction after partial reinforcement (**5**).

An important factor in Sheffield's generalization interpretation is the role of stimulation arising from food particles in the mouth, present on trials immediately following reinforced trials but not on any extinction trial. We have assumed that the degree of persistence of such mouth cues depends upon the kind of reinforcing agent employed. With this in mind, two repetitions of the Sheffield experiment were undertaken, varying only the kind of reinforcement. In Exp. 1, dry food was substituted for the wet mash used by Sheffield as reinforcement; this was expected to increase the persistence of traces and enhance—at least not reduce—the Sheffield effect. In Exp. 2, water was used as the reinforcing agent in an attempt to minimize the duration of the reinforcement stimulus trace since, presumably, water does not persist as a mouth cue as long as food. Therefore, following Sheffield's reasoning, not only

[1] This article is based on the master's theses of W. W. and E. J. W. (**12, 13**) which were directed by A. A. The authors are indebted to Dr. E. Lee Hoffman for his advice and help in the statistical analysis.

[2] Now at Harvard University.

should there be no difference in resistance to extinction between partial- and continuous-reinforcement groups after spaced training, but little or no difference even after massed training with water as the reinforcing agent, since reducing the duration of the reinforcement traces is, in this respect, the equivalent of increasing the time between trials.

These experiments also provide data pertinent to Sheffield's findings concerning massed versus spaced extinction (**8**). She found greater resistance to extinction with massed than with spaced extinction trials, particularly after massed acquisition.

Method

Subjects.—The *S*s were 144 naive rats, 72 in each experiment. These included male and female, hooded and albino, and ranged in age from 4 to 10 mo. For each of nine replications, however, the strain, sex, and age of the *S*s were held as constant as possible within and between the two experiments. Twenty-two additional *S*s were discarded during pretraining trials for failure to meet the criteria for service as *S*s.

Apparatus and procedure.—Except for differences in motivational-reward conditions, the two experiments duplicated as exactly as possible Sheffield's experimental apparatus, design, and procedure. Two identical straight-alley runways were used, constructed in every respect according to Sheffield's specifications (**7**). Each runway was illuminated by a 40-w. fluorescent tube about 4 ft. above the runway and in line with it.

The experiment was run in three stages: 10 pretraining trials on Day 1, 30 acquisition trials on Day 2, and 30 extinction trials on Day 3. The three main variables were intertrial interval in acquisition (15 sec. vs. 15 min. or M vs. S) reinforcement ratio in acquisition (50% vs. 100% or P vs. C) and intertrial interval in extinction (15 sec. vs. 15 min.). This required four acquisition groups (MP, MC, SP, SC) and eight extinction groups (MPM, MPS, MCM, MCS, SPM, SPS, SCM, SCS) for complete counterbalancing. Nine replications were run in each experiment and eight *S*s were used in each replication, one *S* being run under each of the eight extinction conditions. Each of the two *E*s ran one-half of the *S*s in each condition of each of the two experiments.

During a 2-week taming period, *S*s were handled and adjusted to the drive regimen of the experiment. The *S*s in Exp. 1 were reduced to 10 gm. of food per day until the day before the pretraining trials, when their ration was cut to 6 gm. for the 3-day experimental period. The *S*s in Exp. 2 remained on a 16½-hr. water deprivation schedule from the first day of the taming period until the end of the experimental period.

A feature of all experiments involving a comparison of massed and spaced practice—and accentuated in the present design—is that the massed group receives the same number of reinforcements as the spaced group in a much shorter space of time. According to the present design, the massed *S*s receive 30 reinforcements on Day 2 in about 15–20 min., while the spaced *S*s require about 7 hr. to complete 30 trials. We would expect, therefore, the drive level on successive trials to be higher in the spaced group than in the massed group, making questionable the validity of any comparison between spaced and massed acquisition. While acquisition comparisons are not, per se, an important aspect of these experiments, we attempted to offset this factor somewhat by deviating from the Sheffield procedure in the following ways: (*a*) All *S*s in each experiment were brought into the experimental room on the acquisition and extinction days (Days 2 and 3) under the same conditions of deprivation. This meant that no *S* was given the balance of its daily ration on Day 1 or Day 2 until all *S*s had completed trials for that day. (*b*) The running of massed *S*s was stretched out over the 7-hr. period required for the spaced acquisition trials, so that some massed trials were run with the earlier spaced trials and others with the later spaced trials. For any given replication, one *E* ran all massed and the other all spaced acquisition trials; but which *E* ran the massed trials changed from one replication to the next. Since each *E* ran the same *S*s on both acquisition and extinction, and would therefore be running both massed and spaced *S*s at extinction, one-half of the massed *S*s were run before and one-half after all of the spaced *S*s on Day 3.

As was indicated earlier, each of the experiments employed a reinforcing agent different from Sheffield, who used a wet mash of dog chow and flour. In Exp. 1, dry dog-chow food pellets were used; 30 such pellets, given as reinforcement on the acquisition day, weighed 2 gm. The reinforcement in Exp. 2 was two drops of water at room temperature; the 60 drops which reinforced acquisition measured about 4 ml.

After all *S*s had completed their trials on the pretraining and acquisition days, food deprivation groups received the remainder of the 6-gm. daily ration, and water deprivation groups had access to water for 45 min.

Response time, the dependent variable, was measured from the raising of the start-box door to entry into the goal box.

Statistics.—Sheffield reports time scores in acquisition and frequency scores in extinction. Here, frequency scores are reported for both acquisition and extinction, although response-time medians are also reported in Tables 1 and 3. In every case, values in the graphs and in the statistical analyses are numbers of responses at or below the over-all median response time for the appropriate series. The decision to employ these frequency measures was based on three considerations: (*a*) to have the same type of values in all graphs and all statistical analyses; (*b*) to make the results more directly comparable to Sheffield's; and (*c*) to obtain values which could be subjected to analysis of variance and covariance—the time scores deviated from normality and homogeneity of variance even after a square-root transformation.

Instead of the *t* tests used by Sheffield, the acquisition data were subjected to analysis of variance, and extinction data to analysis of variance and covariance. The results of both analyses are offered particularly in regard to the partial reinforcement variable. The analysis of variance of extinction data would test the notion that partial reinforcement leads to greater resistance to extinction *despite* possibly a lower level of acquisition performance in acquisition; the analysis of covariance tests resistance to extinction following partial and continuous reinforcement adjusting for any inequality in acquisition performance.

Results [3]

Experiment 1

This repetition of the Sheffield experiment used dry food as reinforcement instead of wet mash.

Acquisition.—Since we were interested in the final level of acquisition performance, an analysis of variance was performed to test the equality of the four groups on the last 15 of the 30 acquisition trials. The score for each *S* in this analysis was its number of responses at or below the over-all median response time. The total of such responses for each group along with the group median response times on the last 15 trials are given in Table 1. The analysis indicated that variation due to differences among replications was significant beyond the .01 level. With this factor isolated, the difference between massed and spaced training groups fell just short of significance: F was 4.02, significance at the .05 level required an F of 4.14. Continuous reinforcement yields higher final acquisition values than partial reinforcement, but the difference is short of significance at the .05 level ($F = 2.73$). No interaction variance approaches significance at the .05 level. The picture presented by these acquisition data is similar to that of the Sheffield data. The final response-time level was about the same in both experiments.

TABLE 1

Median Response Time in Seconds for Last 15 Acquisition Trials and 30 Extinction Trials and Number of Responses at or Below the Over-all Median: Experiment 1

Group	Acquisition		Extinction	
	Median Response Time	Responses At or Below Median	Median Response Time	Responses At or Below Median
MPS	3	184	65	98
MPM			21	168
MCS	3	167	80	88
MCM			38	148
SPS	3	175	22	172
SPM			41	141
SCS	2	235	42	133
SCM			42	138

Extinction.—In plotting the curves of Fig. 1 and in the statistical analyses the response index employed was the frequency of scores at or below the median response time for the 30 extinction trials. Table 1 presents the median response-time values for each of eight extinction groups and number of responses at or below the over-all median for each group. Figure 1 shows both the massed and the spaced partial reinforcement groups more re-

[3] For other graphic and tabular presentations and analyses of the data of Exp. 1 and 2, see the Wilson (13) and the Weiss (12) theses.

sistant to extinction than the comparable continuous reinforcement groups. If anything, the difference seems larger after spaced training than after massed training. The analysis of variance of extinction data, summarized in Table 2, showed only two significant sources of variation, both

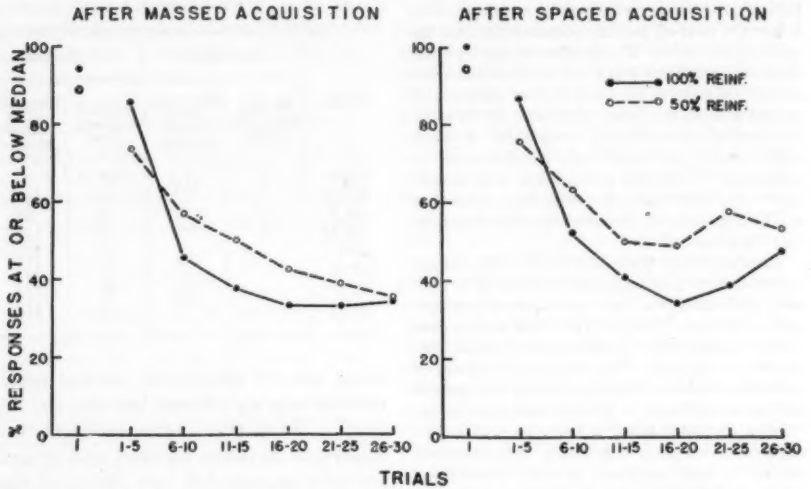

FIG. 1. Extinction curves for food-reinforcement groups after massed and spaced training, showing percentage of responses in 5-trial blocks at or below the over-all median response time. Data of Trial 1 are plotted separately.

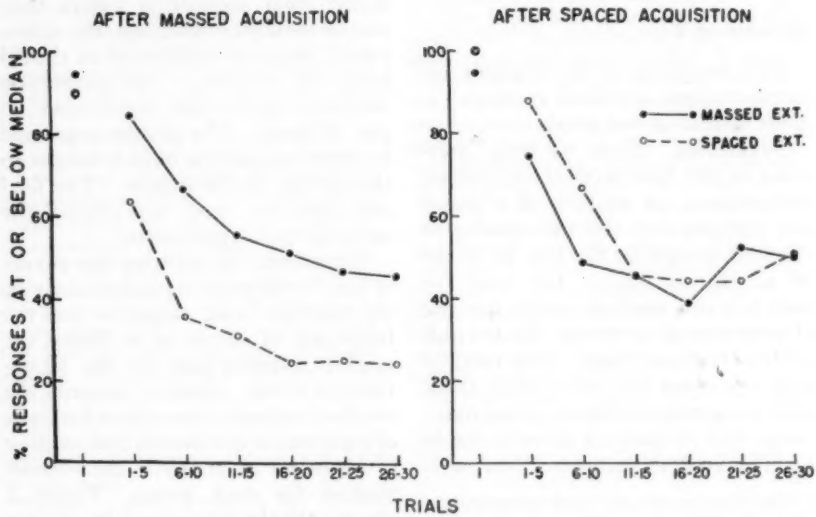

FIG. 2. Extinction curves for food-reinforcement groups showing massed and spaced extinction performance after massed and spaced acquisition. Data of Trial 1 are plotted separately.

at the .01 level—one was for replications, as expected, and the other was a significant interaction between massed vs. spaced acquisition and massed vs. spaced extinction. This significant interaction resulted from the fact, shown graphically in Fig. 2, that the extinction performance of the massed groups was higher than that of the spaced groups after massed acquisition, but at about the same level as the spaced groups after spaced acquisition.

When extinction data were adjusted for differences in acquisition in the analysis of covariance (Table 2), partial reinforcement groups were shown to be significantly more resistant to extinction than were continuous reinforcement groups *regardless of the intertrial interval in acquisition.* The interaction (A × B) between massed and spaced acquisition and partial and continuous reinforcement was small and not statistically significant. If the difference between partial and continuous reinforcement groups in extinction depended upon the distribution of acquisition trials, this interaction would have been large. The analysis of covariance also indicated that massed extinction groups were significantly more resistant to extinction than spaced ($p < .05$). This difference is, however, complicated by the still-present significant interaction (A × C) between the massed-spaced variable in acquisition and the massed-spaced variable in extinction. As in the analysis of variance, variation due to replications was significant.

Experiment 2

This was a repetition of the Sheffield experiment using water as reinforcement instead of wet mash.

Acquisition.—The statistical analysis was based on the number of responses in the last 15 acquisition trials for each *S* at or below over-all median response time. The values shown in Table 3 are, for each acquisition group, the median response time and the number of responses at or below the median response time for the last 15 trials. It will be noted that the number at or below the median was 335 for the continuous groups and 220 for the partial reinforcement groups.

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE AND COVARIANCE: EXPERIMENT 1

Source of Variation	*df*	Analysis of Variance		Analysis of Covariance	
		Mean Square	*F*	Adjusted Mean Square	*F*
A. Massed and spaced acquisition	1	93.36	2.25	25.32	....
B. 100% and 50% reinforcement	1	72.00	1.74	221.84	8.19**
C. Massed and spaced extinction	1	150.22	3.62	137.22	5.60*
D. Replications	8	133.06	3.21**	83.21	3.07*
A × B	1	2.00	....	33.85	1.25
A × C	1	338.01	8.15**	168.42	6.22*
A × D	8	52.01	1.25	54.43	2.01
B × C	1	9.39	....	16.02	....
B × D	8	56.12	1.35	30.93	1.14
C × D	8	35.28	....	21.04	....
Residual†	33	41.47	....	27.07	....

* Significant at .05 level.
** Significant at .01 level.
† Residual *df* are 32 for the analysis of covariance.

TABLE 3

MEDIAN RESPONSE TIME IN SECONDS FOR LAST 15 ACQUISITION TRIALS AND 30 EXTINCTION TRIALS AND NUMBER OF RESPONSES AT OR BELOW THE OVER-ALL MEDIAN: EXPERIMENT 2

Group	Acquisition		Extinction	
	Median Response Time	Responses At or Below Median	Median Response Time	Responses At or Below Median
MPS	7	109	117	119
MPM			50	183
MCS	3	161	120	105
MCM			88	140
SPS	7	111	91	140
SPM			87	149
SCS	3	174	120	89
SCM			78	156

The analysis of variance shows that the higher level of performance attained by the continuous reinforcement groups is significant beyond the .05 level ($F = 6.98$). Distribution of trials had no significant effect on acquisition. The variation due to differences among replications was significant beyond the .05 level. No interaction variance approached statistical significance in acquisition. The final level of performance was lower (higher running time) in this experiment than in Exp. 1 or in Sheffield's study.

Extinction.—The statistical analysis of the extinction data and the extinction curves are based on the number of responses at or below the over-all extinction median. Table 3 gives the number of such responses for each of the eight extinction groups along with the median response time of each group. Extinction curves as a function of acquisition conditions are shown in Fig. 3. After either spaced or massed training, the groups whose acquisition was with partial reinforcement were more resistant to extinction than those which had been continuously reinforced. The comparisons plotted in Fig. 3 do not suggest that the difference in extinction between

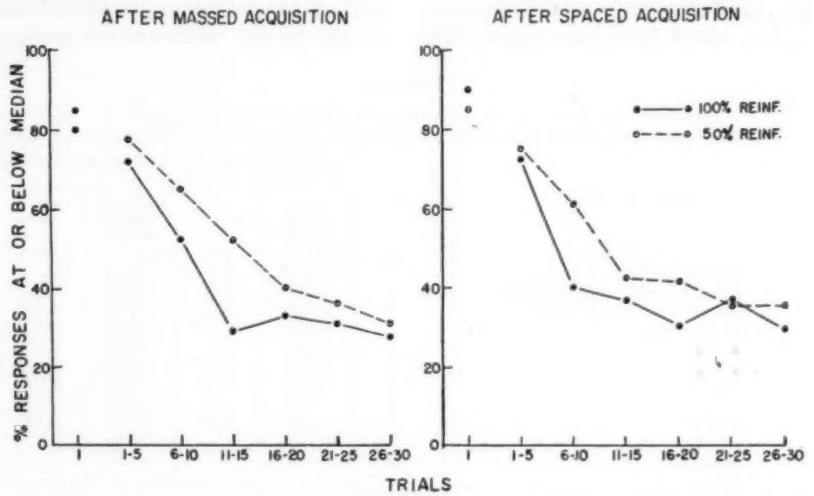

FIG. 3. Extinction curves for water-reinforcement groups after massed and spaced training, showing percentage of responses in 5-trial blocks at or below the over-all median response time. Data of Trial 1 are plotted separately.

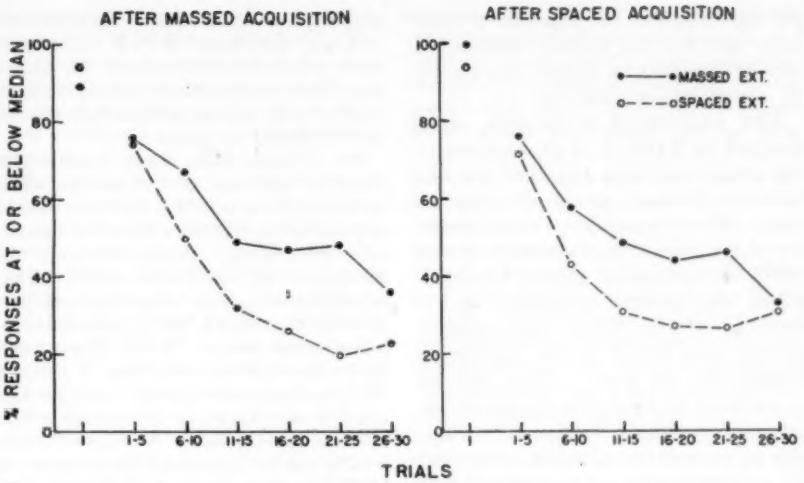

FIG. 4. Extinction curves for water-reinforcement groups showing massed and spaced extinction performance after massed and spaced acquisition. Data of Trial 1 are plotted separately.

partial and continuous reinforcement groups depends upon the distribution of acquisition trials. The analysis of variance of extinction data (Table 4) shows two significant sources of variation, both at the .05 level. As in Exp. 1, one is for replications, as expected; the second, in this experiment, is for the massed-spaced extinction variable. Figure 4 shows graphically the large difference in the curves for massed and spaced extinction after massed or spaced acquisition. The interaction between massed vs. spaced acquisition and massed vs. spaced extinction was very small in this experiment. Variation due to partial as against continuous reinforcement was

TABLE 4

SUMMARY OF ANALYSIS OF VARIANCE AND COVARIANCE EXPERIMENT 2

Source of Variation	*df*	Analysis of Variance		Analysis of Covariance	
		Mean Square	*F*	Adjusted Mean Square	*F*
A. Massed and spaced acquisition	1	2.35	....	.11	....
B. 100% and 50% reinforcement	1	141.68	2.32	413.93	8.88**
C. Massed and spaced extinction	1	425.35	6.98*	541.72	11.63**
D. Replications	8	155.51	2.55*	98.28	2.11
A × B	1	2.35	....	8.72	....
A × C	1	7.35	....	18.42	....
A × D	8	23.91	....	9.68	....
B × C	1	11.68	....	67.93	1.45
B × D	8	81.99	1.35	37.97	....
C × D	8	13.28	....	10.65	....
Residual†	33	60.93	....	46.57	....

* Significant at .05 level.
** Significant at .01 level.
† Residual *df* are 32 for the analysis of covariance.

not significant in the analysis of variance, nor was the critical interaction between massed vs. spaced and partial vs. continuous acquisition.

The analysis of covariance, summarized in Table 4, is an analysis of the extinction data adjusted for differences between groups in acquisition. When an adjustment was made for the superior performance of the 100% reinforcement groups in acquisition, the greater resistance to extinction of the 50% reinforcement groups was found to be significant beyond the .01 level. (The analysis of variance had failed to indicate significance at the .05 level.) The analysis of covariance also demonstrated the greater resistance to extinction of the massed extinction over the spaced extinction groups ($p < .01$). No other variation and no interaction variance was statistically significant.

Discussion

Partial reinforcement and resistance to extinction.—In both experiments—with dry food or water as reinforcement—partial reinforcement groups were more resistant to extinction than continuous reinforcement groups, whether acquisition trials had been massed or spaced. This finding was not significant in the analysis of variance in either experiment, but was in the analysis of covariance. Sheffield's emphasis on the stimulus function of rapidly dissipating traces of reinforcement and nonreinforcement in the explanation of resistance to extinction following partial reinforcement is not supported in either case. To support Sheffield's interpretation it would have to be shown that massing of acquisition trials enhanced the resistance to extinction of the partial reinforcement relative to the continuous reinforcement group, particularly in Exp. 1, and that spacing of acquisition trials prevented any difference between partial and continuous groups in extinction, especially in Exp. 2. The fact that in all four extinction analyses—analysis of variance and covariance for each experiment—there was not a single significant A × B interaction (interaction between massed vs. spaced acquisition and partial vs. continuous reinforcement) argues convincingly against the Sheffield hypothesis.

It is conceivable that carried-over traces of previous goal events do affect resistance to extinction, but they would appear to be, at best, a factor of secondary importance. Crum, Brown, and Bitterman (2) found that a 50% delay-reinforcement group was more resistant to extinction that a 100% immediate reinforcement group. Tyler, Wortz, and Bitterman (10) reported that a random 50%-reinforcement group extinguished more slowly than an alternating 50%-reinforcement group. Neither of these results can be accounted for in terms of Sheffield's trace factor. Weinstock (11) spaced partial and continuous reinforcement trials at 24-hr. intervals, precluding any carry-over effect, and the advantage of partial reinforcement groups was still evident.

Slower extinction following partial reinforcement, *whether massed or spaced*, can be accounted for if a learned (associative) factor is postulated as the primary one. Such a factor must be relatively uninfluenced by the kind of reinforcement (e.g., food or water) or the distribution of trials and its effects must increase as acquisition progresses. That persistence of responding, despite nonreinforcements, develops as partial reinforcement training proceeds is recognized implicitly in the Sheffield procedure. On the first day of the experiment, 10 successive reinforced trials are run. On the second (training) day, the maximum number of consecutive nonreinforced trials is four, and these are reserved for the latter half of training "in order to prevent partial-reinforcement animals from refusing to run in the early part of training" (7, p. 516). It seems obvious, then, that by the time the extinction series is begun the partially reinforced *S*s have been trained to persist in running despite nonreinforcements.

The mechanism through which this persistence is developed can be conceived of as the conditioning of running to re-

sponse-produced stimuli related to anticipatory frustration ($r_F \rightarrow s_F$). By this reasoning, runway stimuli are first conditioned, in the acquisition series of the partial-reinforcement groups, to fractional anticipatory goal (reinforcement) responses, $r_R \rightarrow s_R$. Then, when nonreinforcement comes to be frustrating,[4] the same runway stimuli are conditioned to the frustration reaction and $r_F \rightarrow s_F$ results. Although s_F would tend initially to evoke responses antagonistic to running, as long as running continues in partial reinforcement training s_F should become conditioned to the running response. If continued running is conditioned to stimuli produced by anticipatory frustration responses, as well as anticipatory reward responses, the partial groups would be more resistant to extinction than the continuous reinforcement groups. These implicit anticipatory conditioned mechanisms would develop under either massed or spaced acquisition conditions. Sheffield (**7**, p. 521) recognizes the possibility that some such mechanism might operate "in addition to cues from the frustration which carries over from a nonreinforced trial to the start of the following trial with massed training." Our results, and others cited already, suggest that the anticipatory frustration factor may be the more important of the two determining increased resistance to extinction following partial reinforcement.

Distribution of extinction trials and resistance to extinction.—Sheffield (**8**) reported greater resistance to extinction with massed than with spaced extinction conditions. Stanley (**9**) has confirmed Sheffield's findings for a response time measure—he terms this "a vigor, and presumably a drive, measure of performance"—but has shown that an association measure of performance (percentage of correct runs) shows spaced extinction groups more resistant to extinction. In Sheffield's study and in the comparable portion of Stanley's, massed groups were significantly more resistant to extinction than spaced groups after massed acquisition; the extinction difference is in the same direction, but is not significant after spaced acquisition. This suggests that two factors are operating: (*a*) a motivational effect related to nonreward (frustration) which would show up particularly under massed conditions (**1, 8, 9**), and (*b*) generalization decrement due to shifting of intertrial intervals from acquisition to extinction. According to this reasoning the massed acquisition-massed extinction groups should be most highly resistant to extinction, and the massed-spaced groups should extinguish most readily.

Our repetitions of this portion of the Sheffield analysis showed Exp. 2 in complete agreement with her finding and Stanley's; the results of Exp. 1 were in partial agreement. In Exp. 2 massed extinction groups were more resistant to extinction than spaced groups after spaced or massed training, although the difference was somewhat larger after massed acquisition. The interaction between massed-spaced acquisition and massed and spaced extinction was not significant. In Exp. 1, the massed-massed group resisted extinction significantly more than the massed-spaced group, but the difference between spaced-massed and spaced-spaced groups was slightly (not significantly) in the reverse direction. This, by itself, would not rule out the two-factor interpretation; however, in Exp. 1 there was no significant difference in resistance to extinction between spaced-spaced and massed-massed conditions, neither of which was subject to generalization decrement. This aspect of the results of Exp. 1 is contrary to the interpretation offered above.

Possibly the disagreement in this aspect of the results of our two experiments can be accounted for in terms of the general performance level attained by the food- and the water-reinforcement groups. The food-deprivation *S*s showed consistently lower response times than

[4] A study by Roussel (**6**) and other unpublished data from our laboratory support the notion that nonreward is not frustrating early in partial (50%) reinforcement training, but becomes more frustrating as training progresses; also, there is an indication that the degree of the frustration effect is a function of the amount of "reward expectation" which has already developed.

the water-deprivation *Ss* in acquisition and extinction (compare Tables 1 and 3). If we take this to mean that the food-deprivation *Ss* were nearer to their limit of performance, it could be argued that the frustrative effect in extinction could show up more in Exp. 2 than in Exp. 1, and the generalization decrement more in Exp. 1 than in Exp. 2. Were this the case, the reversal in Exp. 1 between the spaced-massed and spaced-spaced groups —the only reversal in either experiment —could be attributed to a generalization decrement in switching from spaced acquisition to massed extinction which more than offset the facilitating frustration effect of massing extinction trials.

Summary

Two experiments were performed to test Sheffield's interpretation of resistance to extinction following partial reinforcement. Sheffield's design and procedure were duplicated and the experiments differed from hers only in the motivational-reward conditions employed. In Exp. 1 hungry *Ss* were rewarded with dry food, while in Exp. 2 thirsty *Ss* were rewarded with water. Sheffield had employed a wet mash as reinforcement.

The results were not in agreement with Sheffield's, and did not support her interpretation. Partial (50%) reinforcement groups were more resistant to extinction than continuous (100%) groups regardless of intertrial interval in acquisition or kind of reinforcement. These differences were significant only in an analysis of covariance which adjusted the groups for differences in acquisition levels of performance. An associative mechanism was postulated to account for the results.

Although, in Exp. 1, massing of trials increased resistance to extinction, the effect of distribution of extinction trials on resistance to extinction depended also upon the distribution of acquisition trials: the switched groups were less resistant to extinction than the unswitched ones. In Exp. 2, massed extinction groups were more resistant to extinction than spaced ones after both massed and spaced training and after both partial and continuous reinforcement, although here too generalization decrement seemed to be a factor.

References

1. Amsel, A., & Roussel, J. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.
2. Crum, J., Brown, W. L., & Bitterman, M. E. The effect of partial and delayed reinforcement on resistance to extinction. *Amer. J. Psychol.*, 1951, **64**, 228–237.
3. Grant, D. A., Hornseth, J. P., & Hake, H. W. The influence of the intertrial interval on the Humphreys' "random reinforcement" effect during the extinction of a verbal response. *J. exp. Psychol.*, 1950, **40**, 609–612.
4. Grant, D. A., Schipper, L. M., & Ross, B. M. Effect of intertrial interval during acquisition on extinction of the conditioned eyelid response following partial reinforcement. *J. exp. Psychol.*, 1952, **44**, 203–210.
5. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
6. Roussel, J. Frustration effect as a function of repeated nonreinforcements and as a function of the consistency of reinforcement prior to the introduction of nonreinforcement. Unpublished master's thesis, Tulane Univ., 1952.
7. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.
8. Sheffield, V. F. Resistance to extinction as a function of the distribution of extinction trials. *J. exp. Psychol.*, 1950, **40**, 305–313.
9. Stanley, W. C. Extinction as a function of the spacing of extinction trials. *J. exp. Psychol.*, 1952, **43**, 249–260.
10. Tyler, D. W., Wortz, E. C., & Bitterman, M. E. The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *Amer. J. Psychol.*, 1953, **66**, 57–65.
11. Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24 hour intertrial interval. Unpublished doctor's dissertation, Indiana Univ., 1953.
12. Weiss, E. J. Extinction as a function of water reinforcement and distribution of practice. Unpublished master's thesis, Tulane Univ., 1954.
13. Wilson, W. Extinction as a function of food reinforcement and distribution of practice. Unpublished master's thesis, Tulane Univ., 1954.

(Received September 30, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

RESISTANCE TO EXTINCTION AS A FUNCTION OF THE TYPE OF RESPONSE ELICITED BY FRUSTRATION

H. M. ADELMAN AND J. L. MAATSCH[1]

Michigan State College

Most contemporary theories of inhibition, regardless of the type of explanatory concepts involved, explain experimental extinction as a result of the learning of interfering responses. These theories do, however, differ as to the origin of the interfering responses and the principles of learning involved. Hull (**4**) and Miller and Dollard (**6**), for example, posit that interfering responses arise from the accumulation of intrinsic inhibition resulting from sheer performance of the learned response. These interfering responses are variously described as resting responses, relaxation responses, response of not responding, etc., and result in the reduction of the accumulated inhibitory drive state. On the other hand, pure interference theories of extinction as represented by Guthrie (**3**) and others (**1, 8**) maintain that any new response which consistently occurs during extinction will be learned and interfere with the original response, since sheer contiguity of stimulus and response are considered sufficient to produce learning.

The theoretical orientation of the present study is similar to Rohrer's (**7**) in that it holds that the new interfering responses that appear during extinction are elicited by frustration stimulation (s_f) (**5**). Frustration stimulation occurs as a result of the disruption of a learned sequence, e.g., the omission of food where previously present as a reward, blocking an instrumental response, etc.; and s_f elicits a characteristic class of responses (Rs_f).[2] The elicitation of Rs_f by s_f constitutes an instance of reinforcement[3] just as *elicitation* of any other characteristic class of responses by some stimulus or stimulus complex serves as an instance of reinforcement. Examples of reinforcement so defined would include any US and its appropriate UR class, e.g., the elicitation of "salivation" by meat powder, "avoidance" by shock, or "approach and eat" by food. If an instance of reinforcement occurs, then a habit involving the *total* stimulus complex and response class elicited will occur and/or be strengthened. Therefore, the repetition of nonreward trials in a previously rewarded situation will result in: (*a*) the elicitation of new responses (Rs_f), (*b*) the strengthening of a habit of so responding to the cues of the previously rewarded situation, and (*c*) the interaction or interference of the new habit with the old habit ($_sH_R$).

Since s_f may elicit a series of different UR's, e.g., preening, defecation, exploration, etc., the response tendency

[1] The authors wish to express their gratitude to Professor M. Ray Denny for making the present study possible and for assistance in preparation of the manuscript. The second author is now at the South Dakota State College.

[2] Frustration is defined as follows: Frustration stimulation (s_f) will occur in a learned sequence whenever the elicitation of a learned response (*CR*) results in the occurrence of a stimulation complex (*S*) interrupting performance of the learned sequence. $_sr_f$ will elicit members of a characteristic class of responses (Rs_f). (Notice the first sentence of the definition suggests the operational reduction basis of this construct, while the second sentence defines the theoretical function of the construct once asserted.)

[3] A reinforcement is defined as follows: A reinforcement (*E*) will occur whenever there occurs a stimulus (s) or a stimulus complex (*S*) that elicits a characteristic response (r). Given the occurrence of a reinforcement, there will result an increment to a tendency ($_sH_R$) for that complex to evoke a member of that response class (*R*).

strengthened will depend upon the type of response which removes *S* from the frustrating situation (5). Thus, for example, the extinction situation may be structured to specify the terminal response to frustration, i.e., instrumental escape response to s_f, or the frustrating situation may be terminated independent of the response to s_f, as in the conventional extinction situation, where time in the goal box serves as the criterion for removing *S* from the situation. In the former case, with continued extinction trials one response is consistently reinforced, while in the latter case no one type of Rs_f is consistently reinforced. As a consequence, interference during extinction, in one case, would involve the competition of the original learning with one consistently reinforced habit based upon s_f. In the other case, since no one habit is consistently reinforced, more extinction trials would be necessary for some one habit based upon s_f to accrue sufficient strength to interfere with the original habit. Therefore, we may expect a more rapid extinction of the original response where the Rs_f is environmentally specified than where it is not.

The *interaction* between the original habit and the habit based upon s_f will depend upon the effect of stimulus generalization in the situation and upon the degree of incompatibility of the two response classes. Stimulus generalization will determine *where* in the situation the two habits will *interact* and the relative strength of each to the common stimulus complex. For example, since the habit based upon s_f is assumed to behave as any other habit, we may expect it to generalize to the same temporally antecedent complexes to which the original habit has generalized, other things being equal. In this case, interaction would occur throughout the originally learned sequence. However, the *observable* product of this interaction will depend upon the types of response classes elicited by the two competing habits. For example, if the two response classes are such that they may occur simultaneously, i.e., they are compatible, no interference will be observed. But if the two response classes cannot occur simultaneously, i.e., they are incompatible, interference or "extinction" will be observed.

On the other hand, if generalization of the habit based upon s_f does not occur or is prevented by a discrimination learned through blocking the Rs_f in all places except where originally elicited, then no interaction or interference with the original habit should occur. In other words, during extinction the elicitation of the original habit would merely lead to s_f, and then and only then would Rs_f occur. The two habits would chain and the original habit would be "fixated" with respect to that extinction situation, that is, unless the terminal Rs_f is blocked creating a new source of s_f. An Rs_f that fails to interact during extinction in such a manner will be called temporally compatible with the original habit.

In general, then, the significant theoretical variables determining extinction are considered to be: (*a*) the strength of the original habit, (*b*) the number of occurrences of s_f, (*c*) the specification or lack of specification of Rs_f, (*d*) the type of Rs_f specified and its relationship to the original response, (*e*) the influence of generalization and discrimination in determining the manner in which the two habits interact.

Specifically this study was designed to test this formulation by comparing the effects of three types of Rs_f upon the rate of extinction of a simple running response in a straight alley maze. An Rs_f of jumping from the goal box was considered a specified, temporally compatible response since it could occur only at the termination of the running response. An Rs_f of recoiling from the goal box into the alley was considered a specified directly incompatible response since an anticipatory occurrence of "avoidance of the goal box" and "approach to the goal box" could not occur simultaneously. The third type of Rs_f was unspecified and presumably incompatible since the conventional extinction procedure of confining *S* in the goal box for 20 sec. was used.

From a consideration of the foregoing, the following hypotheses are demanded: (*a*) the jump group will show little or no

extinction of the original response; (*b*) the recoil group will show relatively rapid extinction of the original response; and (*c*) the normal extinction group will fall somewhere between the remaining two groups with respect to resistance to extinction of the original response.

METHOD

Apparatus.—The apparatus was a conventional straight-alley maze. The starting box section was 11 in. square, painted flat gray, and covered with a piece of ¼-in. clear plate glass. The runway was 18 in. long, 5½ in. wide, 8½ in. high, of natural plywood color, and covered by ¼-in. hardware cloth. The goal box section was 11 in. square and 10 in. high, painted black, and covered with a piece of ¼-in. clear plate glass. A natural plywood guillotine door separated the goal box from the straight alley and a semicircular piece of black Bristol board (11-in. radius) was mounted on top of the goal box on the side facing the alley to prevent *S*s viewing the rest of the maze when perched on top of the goal box. A 2-in. black ledge was attached to the top external part of the goal box on the three remaining sides to facilitate perching after *S* had jumped from the goal box.

Subjects.—The *S*s were 30 experimentally naive female hooded rats, 90–150 days old, from the colony maintained by the psychology animal laboratory at Michigan State College.

Procedure.—The *S*s were handled for seven days prior to introduction into the maze. During this time they were put on a 23-hr. feeding schedule and received an average of 9 gm. of Purina Dog Chow checkers daily at the scheduled training time. Throughout the course of the experiment all *S*s were individually fed 9 gm. 10 min. after the end of the daily run. On Day 8, *S*s were introduced into the maze and allowed free exploration for a 1-hr. period. On Day 9 all *S*s were given three spaced (10-min. intertrial interval) acquisition trials. Four spaced trials were given on Day 10, and six spaced trials per day were given on Days 11–15. The time allowed for eating during acquisition was gradually decreased. During the latter phases of acquisition, a 20-sec. period after entering the goal box was allowed for eating. All *S*s ate the ½-gm. reward pellet within this time interval. Twenty seconds after entering the goal box and securing the reward, *S*s were removed to running cages to await the next trial. Thus, each *S* received a total of 37 spaced acquisition trials prior to extinction. Running times were recorded to the nearest second from the time *S* was placed in the starting box until *S* entered the goal-box section.

On the following day, *S*s were divided into three groups on the basis of their performances on the previous day and given one rewarded warm-up trial to indicate the control of group running times prior to extinction. On the next trial (10-min. intertrial interval) extinction of the running response was begun by removing the reward from the goal box.

The groups were differentiated by the type of response allowed following the frustration resulting from nonreward. For the jump group, the plate glass cover of the goal box was removed to allow access to the ledge on top of the goal box. After jumping, *S* remained on the ledge for 20 sec. before being returned to the individual running cages. If *S* did not jump within the 5-min. period, he was aided in climbing to the top of the goal box, by *E* inserting a hand into the box to serve as a step.

A second group, the normal group, was confined for 20 sec. after entering the goal box and then returned to the running cages. This method is frequently used as an extinction procedure in learning studies in this area.

The third group, the recoil group, was allowed to recoil out of the goal box after frustration by leaving the guillotine door between the alley and goal box open. Upon re-entering the alley, the door was dropped behind him and 20 sec. later *S* was removed from the alley or starting box and placed in a running cage. It should be re-emphasized that the groups were differentiated on the basis of the response which occurred *after* frustration of the original response and that *all Ss were retained in the maze for a period of 20 sec. after the response to frustration had occurred.*

The extinction period covered three days of ten spaced (10-min. intertrial interval) extinction trials per day. If *S* did not enter the goal box

TABLE 1

EXTINCTION OF STRAIGHT ALLEY RESPONSE AS A FUNCTION OF TYPE OF RESPONSE ELICITED BY FRUSTRATION STIMULATION

Extinction Condition	*N*	Mean†	*d* Value	Trials to Extinction Medians
1. Jump	10	5.75	$d_{1,3}=4.6$*	30
2. Normal (20-sec. delay)	10	47.5	$d_{1,3}=5.0$*	26.2
3. Recoil	10	60.30	$d_{2,3}=1.3$‡	16.0

* .01 = 3.40.

† Numbers represent the group mean of the median score for each *S* during the thirty-trial extinction period.

‡ When *d* was calculated on the basis of the last ten trials, this comparison was significant beyond the 1% level of confidence as were all other comparisons. This criterion is a more conventional but a less conservative one.

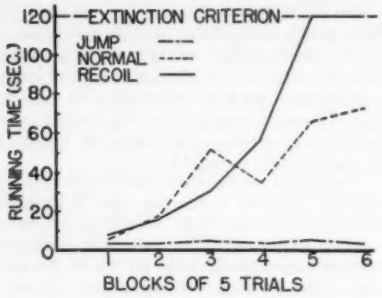

FIG. 1. Group medians of the median running times of each *S* for successive blocks of five trials during extinction.

within 120 sec., it was removed. Two such successive scores were considered as an extinction criterion, even though all *S*s were given all 30 extinction trials. The running times during extinction were recorded in the same manner as employed during acquisition.

RESULTS

The results are presented in Fig. 1 and a statistical analysis of the differences in running times for the 30 extinction trials appears in Table 1. Figure 1 clearly shows that the type of response produced by frustration is a significant variable in the extinction of a simple running response. A response to frustration which opposes or is directly incompatible with the original response (recoil) produces relatively rapid extinction of the original response; and a response which is temporally compatible with the original response (jump) produces little or no effect upon the original response tendency. Normal extinction procedures result in some intermediate effect upon the original response.

Further evidence to support the differences obtained may be found by an analysis of the number of trials to reach the extinction criterion of two successive 120-sec. no-response trials. All ten *S*s of the recoil group, only four *S*s of the normal group, and none of the jump group reached this criterion.

DISCUSSION

The striking differences obtained during extinction of this adient response would seem to indicate that the type of response elicited by frustration and the manner in which it interacts with the original response is an important variable in the extinction process. In general, the results also suggest that the extinction process is one of competition between the original response tendency and the new response tendency which is specified and reinforced through continued frustration of the original response.

An alternative explanation might be that the differences obtained were due to a differential delay in secondary reinforcement through return to running cages. However, the use of a constant 20-sec. interval between the occurrence of the response to frustration (R_{Sf}) and return to the running cages controlled the effects of this variable. Furthermore, since both the jump and recoil groups were required to learn specific responses with respect to the goal box, both groups remained in the frustrating situation for longer periods of time in the initial stages of extinction than did the normal group. However, in the latter stages of extinction, since learning of the response to frustration progressed with continued extinction trials for both of these groups, they spent less time in the goal box as compared with the normal group. Since these different treatments produced similar experience with the goal box during extinction, but lead to radically different results, the postulation of differential effects of secondary reinforcement produced, presumably, by experience with the goal box receives no support.

Some incidental findings lend further support to the theoretical position previously elaborated. For example, jump *S*s tended, during extinction, to make anticipatory jump responses immediately upon being placed in the starting box. Had these anticipatory jump responses been allowed, very rapid extinction of the original response would have re-

sulted. Since this was predicted prior to the study, steps were taken to prevent anticipatory occurrence of the jump response. The starting box was made taller than the goal box with no ledge, and was differentially painted. However, these steps were only partially successful in preventing generalization of the jump response to the starting box. As a consequence, a glass cover had to be utilized during extinction trials. Thus the experimental situation assured the occurrence of the jump response only at the termination of the original habit. Anticipatory jump behavior was not observed in the other groups. This would tend to support the prediction that the type of Rs_f elicited by s_f is the type of response which through generalization inhibits or facilitates the original habit.

From the present theoretical position, extinction data for the recoil group may also serve as a learning curve of avoidance of a frustrating state of affairs (s_f) in the face of a strong competing approach response. This relatively rapid learning is in accord with data soon to be published which demonstrates, among other things, that the learning to escape (jump) a nonreward situation (s_f) where there is no competition from the original response, proceeds very rapidly, faster in fact, than learning through the use of a food reward for hungry *S*s not previously frustrated.

It is entirely possible that these heretofore uncontrolled variables may be instrumental in producing the relatively great variability found in extinction data. The usual extinction procedure is one of retaining *S* in the goal box for a specified time irrespective of the response produced by frustration. In view of the present study it would seem advisable to terminate the extinction trial upon occurrence of a specific response to the frustrating state of affairs, e.g., face washing, recoil, etc., rather than upon passage of a predetermined time.

Summary

The present study was designed to test the hypothesis that the type of response elicited by the frustrating omission of reward and the manner in which it interacts with the original habit are significant determinants of the rate of extinction of the original response. Thirty female hooded rats were trained in a straight alley with food as a reward. For extinction, they were divided into three groups according to the type of response to be learned to the frustrating state of affairs (nonreward). During extinction one group learned a temporally compatible escape response of jumping out of the goal box, while a second group learned a directly incompatible response of recoiling from the goal box. A third group was extinguished according to the traditional procedure of confinement in the goal box for a 20-sec. period after frustration.

The results clearly demonstrate that resistance to extinction is a function of the type of response elicited by frustration and the manner in which it interacts with the original habit. A directly incompatible recoil response to frustration produces relatively rapid extinction of an approach response, while a temporally compatible escape response to frustration produces little or no extinction of the original response.

References

1. Culler, E. Recent advances in some concepts of conditioning. *Psychol. Rev.*, 1938, **45**, 134–153.
2. Festinger, L. The significance of difference between means without reference to the frequency distribution function. *Psychometrika*, 1946, **2**, 97–105.
3. Guthrie, E. R. *Psychology of learning.* New York: Harper, 1935.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Maatsch, J. L. Reinforcement and extinction phenomena. *Psychol. Rev.*, 1954, **61**, 111–114.
6. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
7. Rohrer, J. H. A motivational state resulting from non-reward. *J. comp. physiol. Psychol.*, 1949, **42**, 476–485.
8. Wendt, G. R. An interpretation of inhibition of conditioned reflexes as competition between reaction systems. *Psychol. Rev.*, 1936, **43**, 258–281.

(Received August 13, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

EFFECTS OF SEQUENTIAL DEPENDENCIES ON INSTRUMENT-READING PERFORMANCE [1]

VIRGINIA L. SENDERS AND JEROME COHEN

Antioch College

A typical experiment on the principles of dial or instrument design presents *S* with brief exposures of simulated instruments and requires him to make quantitative readings. Such experiments are almost invariably characterized by the extreme care taken by *E* to prevent *S* from predicting in advance anything about the pointer setting which is to be exposed. Successive pointer settings are randomized and independent. While such a procedure makes for clean experimental design and straightforward statistical analysis of results, it also makes the experimental situation very different from the operational situation to which the results are to be applied (**4**).

In the operational situation, the different pointer settings have unequal probabilities of occurrence; sequential dependencies make still more unequal the probabilities of various pointer settings following a previous pointer setting of some known value. A pilot reading an altimeter knows that a value of 15,000 ft. is more probable than one of 90,000, and he knows that if the last reading was 5,000 ft., probabilities of settings between 4,500 and 5,500 ft. are greatly increased over those of, say, 15,000 to 20,000 ft. Another way of saying all this is to assert that input information (from one instrument) in bits, is much less in the operational than in the experimental situation.

In the absence of experimental evidence, we have no way of knowing whether the instrument which permits the most accurate and rapid reading when settings are randomized and stimulus information is high will also prove most satisfactory when various settings are not equally likely or not independent. A program to evaluate the effects on performance of various distributions of absolute and conditional probabilities, and to explore possible interactions between these distributions and instrument design characteristics, therefore seems desirable.

In terms of information, the random distribution of pointer settings which has been typical in past research in the area represents one limiting case. For any given number of alternative pointer settings, information is maximum in this case. Since this limiting case has been thoroughly investigated, we chose for a starting point in our research the other limiting case—that in which stimulus information is zero. This condition occurs when the probability of any specified pointer setting is one. If a watch is broken, its hands are always in the same place and it contains no information. A less trivial case is one in which the proba-

[1] This research was conducted under Contract No. AF 18(600)-50 between the Psychology Branch of the Aero Medical Laboratory, Wright-Patterson Air Force Base, and Antioch College. A fuller discussion and description of results may be found in WADC Technical Report No. 54-253, (**5**), which is available to agencies and contractors of the U. S. Government through the Armed Services Technical Information Agency (ASTIA). The authors wish to acknowledge with gratitude the assistance of Miss Ruth Sherwin and Mrs. Elaine Nierenberg, who gathered the data, and of Mr. Lloyd H. Pounds and Airman Clinton Foulk, who planned and carried out the IBM analysis.

bility of any pointer setting is less than one, but the *conditional* probability of a specified pointer setting, given the previous occurrence of some particular setting, approaches one. No information is contained in the second setting which is not already present in the first, and the second setting is said to be completely redundant. This is the situation in the experiment reported here.

Ten pointer settings were presented with equal frequency; hence, if we ignore temporarily the effects of conditional probability, the amount of information in the set of pointer settings was 3.32 bits ($I_s = \log_2 N$). However, for the experimental groups the ordering of the stimuli was such that, after the first presentation of setting i, the conditional probability of setting i was 1.00. That is, the same setting was presented repeatedly. In a control condition, successive stimuli were made independent so that both the absolute and conditional probability of any stimulus was always approximately .10.

Subjects and Procedure

The *S*s were 140 undergraduates with 20/20 visual acuity in at least one eye. The names of the *S*s were drawn at random from the college directory and were then assigned randomly to five experimental and two control groups.

The *S*s' task was to estimate the position of a pointer on a linear scale projected on a screen and to record that estimate on a mark-sense IBM card.

The stimulus material consisted of 50 lantern slides. Each slide showed a dark linear scale on a light ground; the numbers 0 and 100 marked the ends of the scale, and a simple pointer was positioned above it. Five scales, differing in number of graduation marks, were used. The numbers of graduations between the end marks on the scales were 0, 1, 9, 19, or 99. The same pointer settings were used for all scales. These settings were 7, 15, 23, 34, 41, 58, 69, 76, 80, and 92. Thus the pointer appeared once in each interval of ten units and once on each final digit.

Slides were presented in a standard lantern slide projector from a distance of 25 ft. The image of the linear scale was 32 in. long, which subtends a visual angle of 5°6′ at a viewing distance of 15 ft. The exposure brightness of the light portions of the slide was at a high photopic level. Four timers were used in conjunction with a sliding mirror tachistoscope to give a controlled viewing sequence of 250 ms. pre-exposure dark time, an exposure interval of determined length, an interval of 100 ms. post-exposure dark time, followed by an interval of 11-sec. light time. The timers controlled two overhead "Dim-a-Lights" which gave a dim, diffuse illumination just sufficient for recording results. Five different exposure times were used. These were .100, .225, .550, .775, and 1.000 sec. Each experimental group of 20 *S*s was tested with a different exposure time, one control group was tested with .225 sec., and the other control group with .775 sec. Each group was tested with all five scales, and with all ten pointer positions.

The procedure on any experimental session was as follows: the *S*s entered the room and were seated at tablet armchairs previously arranged for good visibility of the screen. The procedure for marking the IBM cards was fully explained, and *S*s were then told that "this is an experiment to determine the effect of the number of graduations on a scale on the accuracy with which the scale can be read." They were instructed to "estimate the position of the pointer to the nearest integer—that is, the nearest *whole number* between 0 and 99." Each scale to be used (but without a pointer) was exposed first for 10 sec. and then for the exposure time to be used in the experiment proper.

The *S*s in the experimental groups were told, in addition, that once a particular pointer setting had been exposed, the slide would not be changed until ten successive, repeated exposures of it had been given. The *S*s in the control groups were told that the same scale would be exposed ten times in succession but the position of the pointer on the scale would be changed randomly from exposure to exposure.

When all questions had been answered, the experiment began. The lights went out, the first slide was exposed for the specified duration, and after the short postexposure dark period the lights went on. During the 11 sec. that followed *S*s marked their answers and prepared for the next exposure.

For the experimental groups, 10 exposures of the same *scale and pointer setting* constituted a *series*; for the control groups 10 exposures of the same *scale* constituted a *series*. After every 10 series (100 exposures) a rest period of approximately 3 min. was given, and a 10-min. rest period followed the 30th series. An experi-

TABLE 1

EXPERIMENTAL DESIGN

Exposure Number	Experimental Groups*										Control Groups†									
	Series										Series									
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
1	07	92	23	58	80	76	69	15	34	41	23^1	15	92^2	41	92	15	76	15	34	80
2	07	92	23	58	80	76	69	15	34	41	92^1	41	34^1	07	80	69	58	69	58	34
3	07	92	23	58	80	76	69	15	34	41	23^2	41	07^1	23	92	76	15	07	76	76
4	07	92	23	58	80	76	69	15	34	41	15^1	15	15	69	58	23	58	58	07	07
5	07	92	23	58	80	76	69	15	34	41	15^2	76^3	76	15	07	80	69	69	69	23
6	07	92	23	58	80	76	69	15	34	41	41^1	41	80^2	58	34^2	80	34	23	69	07
7	07	92	23	58	80	76	69	15	34	41	69^1	41	80	07	23	80	34	41	41	92
8	07	92	23	58	80	76	69	15	34	41	76^1	23	58^2	07	92	41	92	34	34	80
9	07	92	23	58	80	76	69	15	34	41	15	80^1	69^2	92	34	34	92	69	15	76
10	07	92	23	58	80	76	69	15	34	41	41^2	58^1	07^2	69	58	23	69	23	23	41

* Each *series* consists of ten successive presentations of the same scale position on the same scale. *Exposure i* is considered to be the *i*th presentation of all ten scale positions on the same scale, or the *i*th presentation in each of ten series. A data matrix for a single Exposure number contains 20 responses for each scale position, or a total of 200 responses across the 10 series. A *row* (not a column) in the above table is a single matrix for the information analysis.

† Each series consists of ten successive presentations of the same scale, but not of the same scale position. Exposure *i* is considered to be the *i*th presentation of all ten scale positions on the same scale, regardless of the place in a series where these positions occurred. All scale positions used for Exposures 1 and 2 have been marked with superscripts. A data matrix consists of responses to all scale positions exactly as with the experimental groups.

mental session consisted of 50 series, or 500 exposures, and took approximately 2 hrs. and 40 min.

For both the control and experimental groups, the order in which the slides were presented was subject to several restrictions. Each scale was used twice in every 10 series, and the same pointer position was never used on two succeeding series. Within these restrictions, the order of scales and pointer settings was randomized. The order of the pointer settings was the same for all five experimental groups, but the order of the scales was systematically rotated. For the control groups, the order of the scales was the same as for the two corresponding experimental groups, but the pointer settings within the ten exposures of a series were randomized.

Thus it will be seen that, except for the order in which the stimuli were presented, the treatment of the control groups was identical with that of the experimental groups. They saw the same stimuli the same number of times, under the same conditions of exposure time, illumination, etc. Only the absence of a known systematic ordering (repetition) of stimuli differentiated the control from the experimental condition.

ANALYSIS OF RESULTS

Results were analyzed in terms of Exposure number. To see what this means, let us define two new terms and repeat one previous definition. *An exposure* is the single visual presentation of a pointer setting of a specified scale. It is one flash of light. As previously defined, a *series* consists of 10 successive exposures. For the experimental groups, these 10 successive exposures are always of the same pointer position and scale; for the control groups the 10 exposures of a series are always of the same scale, but the pointer settings are randomly assigned. Thus the amount of stimulus information in a series is zero for the experimental groups, but approaches 3.32 bits for the control groups. The term *Exposure i* is used to designate *i*th exposure of all the 10 pointer positions on a specified scale.[2] For the experimental groups, *Exposure i* refers also to the *i*th member of each of the 10 series used on a scale, but this restriction does not hold for the control groups. For both experimental and control groups, therefore, the amount of stimulus information on any Exposure number is 3.32 bits. For the experimental groups, however, none of the information on Exposures 2 through 10 is new—each pointer setting has occurred in a series of 10 repetitious exposures, and all of the 3.32 bits are therefore redundant. This is not true for the control groups. The experimental design is further clarified by Table 1.

[2] Throughout the text, capitalization of the word Exposure implies that it refers to ten visual presentations rather than one.

The data were analyzed in two different but related ways. The amount of information transmitted on each Exposure was determined for the different experimental conditions by the techniques described by Garner and Hake (1); and a conventional error square analysis was performed for the same conditions.

There are two important objections to the use of an information analysis on such data as these. First, when we use data from pooled *Ss*, we are treating these data as if a group of *Ss* could be thought of as a communication channel. While such a concept is not beyond the bounds of an agile imagination, the results would then apply only to a situation where pooled judgments were to be used. Great caution must therefore be applied in interpreting the group curves; they may not be generalized to the function for a single individual.

Second, information measures are theoretically based on probabilities while in our case they are obtained from relative frequencies with a fairly small *N*. The difficulty caused by this fact is principally in our interpretation of stimulus information. For any scale, ten pointer positions were used, and the stimulus information was therefore 3.32 bits. The *Ss*, however, were not told that only 10 pointer positions would be used, and any previous experience they had had with scales would suggest that there were actually 100 alternatives, or 6.64 bits. An artificial limit is therefore imposed upon the amount of information which can be transmitted in this stimulus situation.

In spite of these objections, information analysis may be extremely useful. If the experimental situation be regarded, not as the conventional input-channel-output system but as a group of individuals giving responses to stimuli, the information measure is seen to result from two components: (*a*) the ability to give different responses to different stimuli, and (*b*) the tendency to give different responses to the same stimuli. It is, therefore, analogous to an *F* ratio, except that it gives less weight to deviant responses.

In addition to the information estimates, variable and constant error square terms were obtained for each combination of exposure time, Exposure number, and scale. The term "mean square error" refers to the sum of the variable and constant mean square errors. The relation between error measures and information measures has been adequately discussed elsewhere (2); suffice it to say here that the two are not independent. In our case the correlation between the logarithm of the mean square variable error and the information transmitted (computed for the experimental groups) was found to be −.86, the anticipated degree and direction of relationship.

Results

Table 2 presents I_t for all the individual experimental and control groups. When the same pointer position is exposed repeatedly, and *Ss* know that this is to be the procedure, there is a marked negatively accelerated increase in the amount of information transmitted on successive Exposures, as may be seen in Fig. 1. Analyses of variance show the variance due to Exposure number to be significant at better than the .001 level for the experimental groups but insignificant for the control groups.

The data in Table 2 may be plotted separately for each combination of scale design and exposure time, and may be combined to give separate plots for each exposure time averaged over scales, and for each scale averaged over all exposure times. These plots show the value of the asymptote and the rate of approach to it to be determined by the interaction of exposure time and scale design. Scales with no graduations, or with only one graduation, never reach as high a maximum as those with more. On the other hand, a large number of graduations does not result in a high asymptote unless the exposure time is quite long.

Results from the error analysis

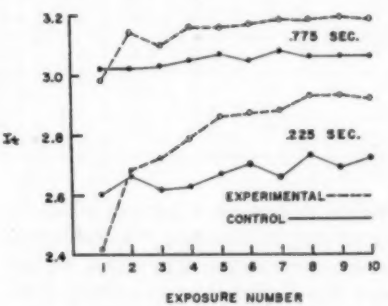

Fig. 1. Information transmitted as a function of Exposure number. Comparison of experimental and control groups.

TABLE 2

INFORMATION TRANSMITTED IN BITS (I_t) AS A FUNCTION OF EXPOSURE NUMBER

Exp. Time (Sec.)	No. Grad.	Exposure Number									
		1	2	3	4	5	6	7	8	9	10
Experimental Groups											
.100	0	2.46	2.53	2.62	2.66	2.74	2.67	2.70	2.72	2.76	2.75
	1	2.51	2.71	2.66	2.66	2.84	2.82	2.86	2.88	2.89	2.88
	10	2.35	2.72	2.82	2.95	2.97	3.02	3.05	3.09	3.09	3.05
	20	2.47	2.70	2.77	2.83	2.86	2.90	2.87	2.97	2.98	2.97
	99	2.39	2.67	2.72	2.86	2.88	2.91	2.93	2.97	2.93	2.97
.225	0	2.48	2.55	2.63	2.61	2.61	2.63	2.68	2.62	2.60	2.61
	1	2.53	2.71	2.78	2.82	2.86	2.79	2.91	2.86	2.76	2.78
	10	2.54	2.71	2.82	2.90	2.93	2.98	2.99	3.00	2.95	3.11
	20	2.56	2.77	2.86	2.87	2.91	2.93	2.89	2.96	2.87	2.96
	99	2.50	2.63	2.80	2.88	2.82	2.94	3.00	3.00	2.93	3.00
.550	0	2.74	2.84	2.86	2.85	2.87	2.91	2.88	2.88	2.86	2.84
	1	2.93	3.10	3.17	3.18	3.16	3.16	3.20	3.14	3.20	3.12
	10	3.08	3.12	3.23	3.28	3.31	3.31	3.31	3.30	3.31	3.26
	20	3.08	3.28	3.29	3.26	3.28	3.32	3.32	3.32	3.29	3.29
	99	2.90	3.09	3.15	3.15	3.19	3.24	3.28	3.28	3.28	3.25
.775	0	2.75	2.96	2.88	2.89	2.86	2.83	2.86	2.90	2.90	2.91
	1	3.01	3.08	3.10	3.14	3.13	3.18	3.18	3.20	3.21	3.16
	10	3.03	3.23	3.24	3.27	3.25	3.28	3.28	3.26	3.28	3.30
	20	3.09	3.22	3.25	3.22	3.26	3.28	3.27	3.24	3.28	3.28
	99	3.00	3.22	3.14	3.26	3.28	3.29	3.29	3.28	3.31	3.27
1.00	0	2.74	2.83	2.92	2.89	2.93	2.95	3.01	3.00	3.02	3.00
	1	2.92	3.02	3.01	3.09	3.13	3.11	3.11	3.09	3.13	3.16
	10	3.29	3.26	3.29	3.29	3.31	3.27	3.25	3.29	3.29	3.27
	20	3.13	3.28	3.29	3.27	3.29	3.30	3.30	3.30	3.30	3.30
	99	3.16	3.27	3.29	3.31	3.32	3.32	3.32	3.32	3.32	3.32
Control Groups											
.225	0	2.48	2.56	2.62	2.67	2.65	2.61	2.82	2.60	2.51	2.65
	1	2.67	2.82	2.73	2.73	2.70	2.71	2.63	2.72	2.76	2.82
	10	2.66	2.65	2.65	2.71	2.78	2.73	2.65	2.84	2.77	2.72
	20	2.67	2.60	2.61	2.54	2.62	2.77	2.61	2.79	2.67	2.76
	99	2.54	2.67	2.46	2.49	2.57	2.66	2.59	2.69	2.75	2.68
.775	0	2.80	2.85	2.76	2.86	2.86	2.84	2.78	2.81	2.74	2.73
	1	2.99	2.98	2.90	2.94	3.06	2.98	3.00	2.91	3.00	2.93
	10	3.07	3.09	3.21	3.17	3.09	3.13	3.23	3.25	3.18	3.21
	20	3.20	3.12	3.14	3.11	3.14	3.20	3.22	3.16	3.20	3.20
	99	3.04	3.09	3.17	3.19	3.18	3.20	3.15	3.19	3.20	3.21

(Table 3) present a picture which is not so easy to interpret. Considering the experimental groups alone, we see from Fig. 2 that there is a sharp drop in mean square error with increasing numbers of Exposures. Figure 3, however, which presents the results for control and experimental groups treated with an exposure time of .225 sec., shows that error drops not only for the experimental group but for the control group as well. Figure 4, pre-

senting comparable results for the groups treated with the longer exposure time, shows error scores for the experimental group to be consistently and increasingly lower than those for the control group, although the absolute magnitude of the difference is not very large. In both cases, the consistency of the trend is more striking for the experimental group, with only one major reversal in the first curve and none in the second.

TABLE 3

MEAN SQUARE ERROR AS A FUNCTION OF EXPOSURE NUMBER

Exp. Time (Sec.)	No. Grad.	Exposure Number									
		1	2	3	4	5	6	7	8	9	10
Experimental Groups											
.100	0	468.2	345.0	265.5	211.6	249.6	170.3	127.8	105.0	141.2	40.5
	1	728.8	262.9	169.8	131.3	157.8	125.5	91.4	75.4	81.4	87.4
	10	673.7	306.6	172.3	108.6	102.9	89.5	64.4	92.0	68.4	89.2
	20	297.2	242.3	131.7	45.1	44.0	22.8	61.1	81.8	48.5	56.3
	99	469.2	282.6	60.1	54.9	30.7	51.4	52.6	109.4	44.1	79.9
.225	0	191.7	82.4	71.4	72.8	88.1	107.5	79.7	114.4	87.7	103.7
	1	237.4	94.1	73.3	88.4	64.5	124.5	84.9	77.8	111.9	78.2
	10	373.6	198.7	223.7	139.4	170.7	139.4	124.6	148.1	129.0	133.3
	20	223.0	100.1	73.2	93.5	65.9	79.9	93.5	54.7	62.2	93.4
	99	187.7	144.3	132.4	124.4	81.1	141.6	80.6	54.2	87.0	101.9
.550	0	20.4	15.0	14.1	14.2	13.7	12.2	12.7	12.7	12.7	17.3
	1	45.9	49.1	5.6	5.1	8.5	6.6	5.4	5.8	6.1	14.1
	10	73.6	17.0	9.1	6.8	8.4	8.4	7.3	7.7	7.4	7.0
	20	95.0	6.8	7.3	6.0	5.0	3.4	2.5	2.7	2.5	1.6
	99	54.8	11.8	10.5	8.8	6.2	3.8	3.8	4.3	3.8	4.3
.775	0	41.4	11.6	11.4	11.9	19.4	21.0	20.1	12.5	12.2	11.5
	1	9.9	7.4	7.1	6.7	5.8	5.0	5.6	4.7	4.9	5.4
	10	28.4	13.3	7.7	8.6	4.4	4.5	3.7	4.4	3.9	3.0
	20	31.8	12.1	6.7	6.0	5.4	5.4	4.4	6.1	4.9	5.2
	99	27.8	12.8	7.2	4.9	5.3	4.5	4.2	3.3	2.9	3.6
1.00	0	17.9	14.3	12.8	12.8	13.8	11.9	11.6	12.3	11.7	13.3
	1	21.9	13.6	17.2	10.1	7.5	7.9	8.4	8.7	8.2	7.5
	10	39.0	4.2	2.0	1.8	2.3	2.1	1.9	2.6	1.7	2.2
	20	18.3	4.5	3.0	3.3	2.6	4.6	3.1	2.1	2.0	2.1
	99	19.0	10.0	7.2	5.6	2.6	2.5	2.1	1.7	1.6	1.2
Control Groups											
.225	0	238.6	190.4	96.2	121.4	34.9	135.8	118.8	170.4	174.2	53.7
	1	214.3	87.8	70.6	163.7	84.7	121.1	236.6	110.9	219.5	85.2
	10	129.5	75.4	122.3	106.1	120.3	44.5	160.3	77.7	120.5	117.7
	20	141.1	45.2	89.8	47.7	229.1	106.1	73.3	56.4	74.3	64.2
	99	140.4	70.6	155.1	117.7	105.9	87.9	87.6	72.1	158.2	83.0
.775	0	17.2	20.4	20.2	14.6	18.7	54.2	53.4	14.3	29.3	29.7
	1	24.9	23.0	57.6	20.1	19.0	19.0	13.1	21.3	19.3	13.7
	10	22.6	11.2	6.8	30.3	20.3	23.3	6.5	6.0	13.8	15.4
	20	34.0	16.6	16.9	14.1	14.1	51.1	10.9	22.0	41.5	13.8
	99	33.0	17.7	14.2	15.5	14.9	19.3	30.0	14.3	7.3	9.2

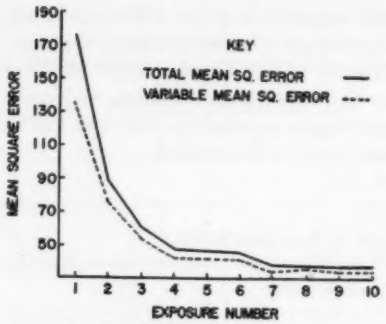

FIG. 2. Mean square error as a function of Exposure number averaged over all scales and times for the experimental groups.

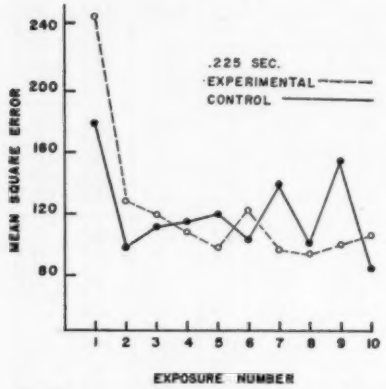

FIG. 3. Mean square error for control and experimental groups treated with an exposure time of .225 sec.

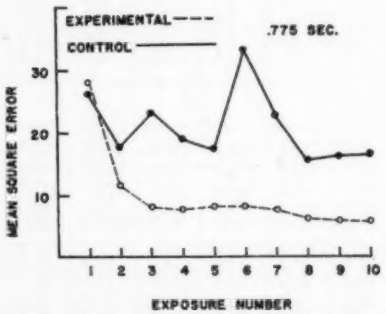

FIG. 4. Mean square error for control and experimental groups treated with an exposure time of .775 sec. Note change of scale from previous figure.

DISCUSSION

These results raise three major questions: (*a*) Why is the difference between experimental and control groups, measured in terms of error, not more clear-cut? (*b*) How could the control groups show no improvement in terms of information, but considerable improvement in terms of error? (*c*) Why is the performance of the experimental groups, however measured, poorer than that of the control groups on Exposure 1?

The answer to the first question can be found by examining the individual curves obtained from Table 3. The experimental group curve for the scale with no graduations, obtained with an exposure time of .225 sec., differs from other experimental group curves by showing a tendency to rise after the second Exposure. The comparable control curve is very irregular but shows an over-all tendency for error to decrease with continuing Exposures. This combination of scale design and exposure time was the only one in which the difference between experimental and control curves was not in the expected direction. If Fig. 3 were redrawn with the ungraduated scale omitted, it would be seen that there is a clear tendency for experimental group error to drop with increasing Exposures, while the control curve shows what appears to be random variability. Referring to the information analysis for this condition, we find that for both groups information shows some increase with increasing Exposures, but in this single case the strength of the relationship appears to be greater for the control than for the experimental groups. Results from the remainder of the experimental groups, whether in terms of information or of error, support the conclusion that *S*s can profit least from the redundancy in the situation when they are working with the ungraduated scale. This conclusion is surprising in the light of the generally accepted principle that redundancy is most valuable when conditions for information transmission are poorest. However, there is evidence to suggest that after the first exposure of a pointer on the un-

graduated scale, individual estimates of pointer position may show increased *internal consistency* without showing greater agreement *between* *Ss*. Our transmission channel is a group of independently operating individuals, and its characteristics are therefore unlike those which would be found in a single individual.

How could control groups show no improvement in terms of information but considerable improvement in terms of error? The answer to this question is to be found in a reconsideration of the actual experimental conditions. The experimental groups had the advantage of redundancy, but the control groups had the advantage of being able to form, earlier in the experiment, an "over-all view of the situation." Consider, for example, the state of the two groups at exposure 23 of any given scale (Table 1). The experimental groups had seen two pointer settings ten times each and one setting three times. The control groups had seen all settings at least once; they had seen five settings twice, one setting three times, one setting four times, and one five times. Thus, at any given Exposure number, it is reasonable to suppose that the control groups, while they were less able to predict *which* setting was to be exposed, were better equipped to describe the population of settings from which it would be drawn. Such knowledge could have resulted in more overlapping of response categories. A setting seen only as "near the high end" (actually 80, let us say) would be more likely to be called 76 or 92 by the control than by the experimental *Ss*, because these *Ss* knew that the latter two categories were among those in the available population of settings. Such responses, if made with sufficient frequency, would have a more marked effect in reducing the estimate of I_t than in increasing the estimate of error.

Figures 1 and 3 raise the question of why the control group performance is superior to that of the experimental groups on Exposure 1. One possible answer, of course, is that there is simply a chance variation in the selection of *Ss* for the groups, the experimental group being poorer. That the experimental group tested with .775 sec. differs very little from its control on Exposure 1 supports this hypothesis. Another possibility is that for some of the experimental *Ss* the first exposure of Series *n* may have been, in effect, exposure 11 of Series *n*-1. That is, they may have overlooked the fact that a new series was about to begin and continued to look at the same point on the scale where the pointer had appeared for the last ten exposures. In the short exposure time given they would not have had time for a shift of fixation when the pointer did not appear as expected. If only first-order sequential dependencies are taken into account, the probability of a *change* of pointer setting is very much less than the probability of a repetition of a setting. Our findings may, therefore, be in accord with the principle that the probable event is more rapidly and accurately perceived than the one which is improbable.

Taken as a whole, the results clearly indicate that performance, measured either in terms of accuracy or in terms of information, is improved by increasing the conditional probability of a pointer setting and thus decreasing the information content of a set of pointer settings. In an operational situation, performance at any given "look" would probably lie somewhere between that of our control and experimental groups, since conditional probabilities would be neither equal to absolute probabilities, nor would they be as high as they were for the experimental *Ss*. The fact that neither of the performance variables we used was a linear function of Exposure number suggests to us the desirability of conducting future experiments on instrument displays in a manner which allows sequential dependencies to affect performance.

In discussing the process of scale reading, Kappauf (3, p. 32) suggests that: "Gross errors of scale use or systematic errors in scale interpretation increase as a function of number of operations the reader has to perform in getting his reading from the scale," and "The adult scale reader does well when his reading can be accomplished in two operations: a number reading operation and some second

operation carried out between numbers, e.g., reading to the nearest marked tenth." It seems likely that our experimental *Ss* may have carried out these operations, and further interpolation operations, in successive exposures, while the control *Ss* were required to carry them all out in one exposure (and were unable to do so accurately). In this respect our experimental *Ss* faced a much more realistic situation than did their controls.

In this experiment we did not find any significant interaction between scale design and Exposure number. However, the case in which the same stimulus is presented repeatedly is not the case in which interactions between display characteristics and Exposure number are most likely to appear. In our next study we hope to use various kinds of nonlinear scales with matched nonlinear progressions of pointer settings. It seems reasonable to expect that in this situation the interaction of Exposure number, kind of progression, and scale design would be highly significant and important.

Summary

An experiment was conducted to determine the effect of sequential dependencies among the pointer settings used as stimuli on scale-reading performance. One hundred *Ss* were presented with ten repeated successive exposures of each of 50 stimuli. Forty control *Ss* received the same stimuli in an unordered sequence. An information analysis and an error analysis were performed on the data.

The results indicate that mere unordered repetition of stimuli results in only a slight, statistically insignificant, increase in information transmitted. Successive exposure of the same stimulus results in a marked negatively accelerated increase in information transmitted by a set of such stimuli, and by a marked decrease in mean square error. The amount of information transmitted is affected both by scale design and exposure time as well as by their interaction. The correlation between the logarithm of the error measures and the information measures was −.86.

Under the conditions of this experiment there was no evidence of an interaction between scale design and Exposure number. The presence of a significant interaction probably depends upon the correspondence between the kind of scale and kind of sequence used.

The implications of these findings for research methodology in the area of instrument and dial design are discussed.

References

1. Garner, W. R., & Hake, H. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.
2. Garner, W. R., & McGill, W. G. Relation between uncertainty, variance, and correlation analyses. *Psychometrika*, in press.
3. Kappauf, W. E. A discussion of scale reading habits. *USAF Tech. Rep.*, 1951, No. 6569.
4. Senders, V. L., & Cohen, J. The influence of methodology on research on instrument displays. *USAF, WADC Tech. Rep.*, 1953, No. 53-93.
5. Senders, V. L., & Cohen, J. The effects of absolute and conditional probability distributions of instrument settings on scale reading: I. Repeated exposures of the same setting. *USAF, WADC Tech. Rep.*, 1954, No. 54-253.

(Received September 25, 1954)

Journal of Experimental Psychology
Vol. 50, No. 1, 1955

VARIATIONS IN SENDING MORSE CODE [1]

CHARLES WINDLE

Human Research Unit No. 3, Ft. Benning, Georgia

In 1897 Bryan and Harter published the first studies of what they called the "telegraphic language" (1). This "language" consists of signals composed of dots, dashes, and spaces in which these signal components are theoretically of prescribed relative durations. According to Bryan and Harter, there are in practice characteristic deviations from the expected dot : dash : space ratios, such deviations being peculiar to particular individuals but consistent for each individual. Deviations among different contextual occurrences of dots, such as the dot in ·— in contrast to that in —·, they called "heterotaxic variations." Deviations among repetitions of a given contextual dot such as the dots in repetitions of ··—, they called "homotaxic variations." Bryan and Harter described heterotaxic variations as "intentional differentiation [corresponding to] inflection in speech" (1, p. 44). Homotaxic variations were assumed to be a measure of the unreliability of sending. Bryan and Harter argued that expert senders were characterized by more heterotaxic and less homotaxic variations than were less efficient senders, but did not present sufficient evidence to justify this conclusion.

The only other study of variations in Morse code component ratios among signals was based on group means for inexperienced *S*s and lacked statistical estimates of reliability (4, 5). Morse code sending performance has received little attention since Bryan and Harter's initial study, in spite of the practical and theoretical implications of the Bryan and Harter suggestions and the need for their validation. The present study proposes to confirm and extend the Bryan and Harter suggestions regarding idiosyncracies in Morse code sending.

METHOD

Subjects.—The *S*s were selected to form criterion ability groups. An inexperienced group was composed of seven male college students completing a course at Columbia University in learning the international Morse code and an assistant for this course without previous sending experience. Each of these eight inexperienced *S*s had less than six months of Morse code sending experience. An expert group consisted of five telegraphers with from 24 to 49 yr. of experience. Although most of the experts were accustomed to sending with the semi-automatic "bug" rather than a key, their sending with a key was found to be competent. Three of the experts used the American Morse code, but the difference in the type of Morse code was thought to be unimportant in the present study.

Apparatus.—A telegraph key was electrically connected to both an audio-oscillator and one channel of a four-channel polygraph. When the key was depressed, the oscillator sounded and simultaneously a visual record was made on tape pulled at 61 mm. per sec. by a constant-speed motor. Measurements were made to the nearest ½ mm., the equivalent of .008 sec.

Procedure.—The *S*s were requested to send cipher material at a "natural" rate. Each was permitted to make his own key adjustment and could set the audio-oscillator to any of three pitches and to whatever volume he preferred.

The cipher material was typed double-space in five-letter groups and was identical for all *S*s. The text was designed to include each of the letters of the alphabet and the single digit numbers four times, these characters being

[1] Part of a doctoral dissertation submitted in partial fulfillment of the requirement for the Ph.D. degree, in the Faculty of Pure Science, Columbia University, in 1952. The author is greatly indebted to Dr. F. S. Keller for his advice and encouragement throughout this study.

distributed so that each would appear once in each quarter of the material.

Since each of the senders made some errors, it was decided to employ only three repetitions of each signal in the computation, eliminating the series with the most errors. Any remaining errors were replaced by correct signals from an adjacent series or, if that was impossible, eliminated from the computation. Signals were discarded as incorrect only in cases of gross inaccuracies such as the omission of a component of the reversal of dots and dashes.

Results

Existence of heterotaxic variation.—The sending performance of each *S* was studied by means of the analysis of variance technique. In these analyses, the heterotaxic variations of Bryan and Harter were estimated in the variance attributable to differences in contextual location. Systematic changes in the dot: dash: space ratios and in speed of sending were controlled by taking out the variance due to replications. The mean square error variance served as an estimate of homotaxic variation or unreliability. Thirty-nine analyses of variance were computed, one for each *S*'s dots, dashes, and intrasignal spaces. In each of these analyses a statistically significant *F* ratio was obtained for the variance within a *S*'s component attributable to signal location ($p < .01$ for 38 analyses; $p < .05$ for one analysis). There is little doubt that heterotaxic variation exists.

Relationship of variations to ability.—Next it was desired to determine whether, as Bryan and Harter claimed, "the accidental variation is a somewhat accurate inverse measure of skill while the variation for inflection is likely to be larger rather than smaller with increasing expertness" (**1**, p. 45). To check on our criterion ability groups chosen on the basis of years of experience, three additional measures were examined: (*a*) number of errors in the four-run trial accepted for each *S*, (*b*) "natural" transmission speed of the signal components (**8, 9, 10**), and (*c*) "natural" transmission speed of the intersignal spaces. Although there was no appreciable difference between groups in number of errors, experts did prove to transmit more rapidly. Chi-square values of cut-off points at 20.00 WPM for tone-speed ($\chi^2 = 9.12$; $p < .01$) and at 11.00 WPM for space speed ($\chi^2 = 5.85$; $p < .02$) are both significant.[2]

Bryan and Harter based their estimates of heterotaxic and homotaxic variations upon the ratio between the mean variation and the mean component length. They could not, however, eliminate the effect of systematic changes in component ratios and in sending speed upon their measure of homotaxic or "accidental" variation. Their measure of heterotaxic variation or "variation for inflection" also was confounded by accidental variation. Estimates and conclusions reached through the multiple classification of variates avoid this confounding.

Crump (**2**) has described a method of estimating variance components in the analysis of variance. Table 1 shows the analysis of variance for the dots of Subject M. S. with both the algebraic terms and the actual values obtained. The best estimate of the average variance contributed by locations (σ_A^2) for M. S. is .19. Similar estimates were made for each *S* for each component.

In each of the six comparisons the mean for the inexperienced *S*s was

[2] Even though Yates' correction was employed, these calculations are subject to criticism for use of small theoretical frequencies. Error scores and space speed were both found to have significantly different variances in the two criterion groups of operators, but not different means according to the more stringent approximation method of testing the significance of differences between means suggested by Cochran and Cox (**3**). The means for tone speed, however, do differ significantly.

TABLE 1

ESTIMATES OF VARIANCE COMPONENTS FROM AN ANALYSIS OF VARIANCE OF OPERATOR M. S.'s DOTS

Source	*df*		Mean Square		Average Mean Square Value	
	Theoretical	Actual	Theoretical	Actual	Theoretical	Actual
Location	$(a-1)$	68	$(M.Sq._A)$	1.10	$(\sigma_{AB}^2+b\sigma_A^2)$	.53+3(.19)
Replication	$(b-1)$	2	$(M.Sq._B)$	.19	$(\sigma_{AB}^2+a\sigma_B^2)$	
Interaction	$(a-1)(b-1)$	136	$(M.Sq._{AB})$	.53	(σ_{AB}^2)	.53

TABLE 2

HETEROTAXIC AND HOMOTAXIC VARIATIONS OF INEXPERIENCED AND EXPERT MORSE CODE OPERATORS

S	Heterotaxic Variation			Homotaxic Variation		
	Dot	Dash	Space	Dot	Dash	Space
Inexperienced						
M. S.	.19	18.20	.18	.53	2.11	.42
C. K.	.16	1.96	.45	.29	1.36	.56
R. W.	.26	17.81	.29	.63	5.88	1.01
C. F.	.64	4.12	.45	.60	3.96	.73
G. T.	3.01	16.90	.92	2.08	3.47	.65
T. K.	.47	5.22	.64	1.09	1.24	.47
L. R.	.21	4.45	.36	.36	3.68	.47
C. W.	.21	15.49	2.70	.82	4.57	2.17
Mean	.64	10.52	.75	.80	3.28	.81
Expert						
F. K.	.15	1.66	.07	.16	.57	.20
R. I.	.15	.21	.17	.28	.39	.15
M. G.†	.06	.18	1.61	.17	.78	.20
L. L.†	.15	2.46	.34	.18	.76	.17
M. F.†	.15	2.67	.16	.14	4.16	.16
Mean	.13	1.44	.47	.19	1.33	.18
t	1.49	3.52*	.68	3.00*	2.14	3.08*

* $p < .05$. Since in each case the variance of the inexperienced operators exceeded that of the experts, the Cochran and Cox approximation method was used (3).
† Operators sending American Morse code.

greater than that for the experts. The finding of Bryan and Harter that homotaxic variation decreases with expertness has confirmation in the present results; their less well-supported claim that heterotaxic variation increases with expertness lacks confirmation.[3]

A structural analysis of heterotaxic variation.—It was hypothesized that

[3] In order to clarify their findings, in part, correlations were run between the rank orders of ability provided by Bryan and Harter and their best measures of homotaxic and heterotaxic variations. The correlation between ability and homotaxic variation (or unreliability) was found to be −.82, significant to the .01 level, while that between ability and heterotaxic variation was +.21, far below the level of significance. It appears that whereas Bryan and Harter had sufficient evidence to claim an inverse relationship between ability and homotaxic variation, they were not justified in claiming any relationship between ability and heterotaxic variation.

TABLE 3

MEAN DURATIONS OF DOTS IN DIFFERENT STRUCTURAL LOCATIONS PRESENTED IN SIGMA SCORES

S	Following Dots			Following Dashes			Initial		
	Prec. dot	Prec. dash	Terminal	Prec. dot	Prec. dash	Terminal	Prec. dot	Prec. dash	Terminal
Inexperienced									
M. S.	−0.10	−0.96	+0.59	−0.15	−1.13	+0.65	+0.96	+0.15	+1.10
C. K.	−0.30	−0.82	−0.02	+0.07	+0.22	−0.50	+0.63	+0.65	+0.91
R. W.	−0.10	−0.40	+0.55	+0.44	+0.44	+1.01	−0.91	−0.55	−0.77
C. F.	+0.83	−0.28	+0.84	−0.14	−1.23	+1.12	−1.11	−1.35	+0.39
G. T.	−0.44	−0.93	+2.34	−0.59	−1.21	+1.08	−0.64	−0.90	+1.37
T. K.	+0.23	+0.45	+0.57	+0.08	−0.64	−0.24	−0.65	−0.56	+1.31
L. R.	−0.07	−0.43	−0.70	+0.22	+1.02	−1.22	+0.67	+1.12	+0.33
C. W.	+0.06	−0.07	+0.18	−0.72	+0.01	−0.43	−0.02	+0.62	+2.04
Expert									
F. K.	+1.52	+0.08	+0.43	0.00	−0.73	−0.18	−0.91	−1.75	−0.96
R. I.	+0.53	+1.36	+0.28	−0.11	+0.13	−0.62	−0.70	−1.02	−0.98
M. G.†	+0.32	+0.73	−0.02	+0.02	−0.05	−1.50	−0.17	+0.56	+0.56
L. L.†	+0.49	−0.79	−0.49	−0.16	−1.49	−1.16	+0.67	−0.09	+0.05
M. F.†	+0.92	−0.41	−0.73	+1.41	+0.05	−1.11	+0.16	−0.49	−0.86

† Operators sending American Morse code.

systematic variations in components' durations might be a function of the types of movements which had just been executed and which were next to be made. Components of signals were classified according to their structural location, that is, according to the types of components which they preceded and followed. Thus, dots were first separated into three groups: those following dots, those following dashes, and those initial in the signal. Each of these groups was again subdivided according to whether the dot in question preceded a dot, preceded a dash, or terminated the signal. Similar classifications were made for the dashes and short intrasignal spaces.[4]

For each *S* the mean duration of each component in each structural location was calculated and subtracted from *S*'s total mean for that component. Each difference was then divided by the *SD* derived from the error variance of the appropriate analysis of variance. All scores, then, were sigma scores indicating relative differences in the durations of particularly located components for each individual. The results of this analysis for one component, the dot, are shown in Table 3.

This structural analysis revealed the following trends:

1. There was a consistency for a given *S* to vary the duration of a component according to the type of component preceded and followed. Thus, when determinations are made for each *S* (e.g., M. S.) of the longest dots of those "following dots" (for M. S., +.59), the location according to element preceded (for M. S., terminal) is unexpectedly often also found to contain the longest dots within the "following dashes" (for M. S., +.65) and "initial" (for M. S., + 1.10) locations ($\chi^2 = 12.81; p < .001$). The same consistency occurs for the dot with respect to the element followed ($\chi^2 = 19.91; p < .001$) and for the dash for the element preceded ($\chi^2 = 50.56; p < .001$), but not for the dash regarding the ele-

[4] For American Morse code signals it was necessary to make a separate classification including long intrasignal spaces.

ment followed or for the short intrasignal space.

2. Within both the dot and space components the dominant feature was the variation in pattern among operators. Seldom was there a consistent trend for *Ss* to have components in a given location different from their means (i.e., zero).

3. As Flik (**4, 5**) claimed, on the basis of less convincing data, the longest dashes were unusually often found to be terminal ($\chi^2 = 11.11$; $p < .001$). Terminal dashes following dots and terminal dashes following dashes tend to be longer than the individual's mean dash more frequently than would be expected by chance ($\chi^2 = 7.69$; $p < .01$; $\chi^2 = 4.92$; $p < .05$, respectively). On the other hand, dashes between dashes and initial dashes preceding dashes tend to be shorter than the mean ($\chi^2 = 7.69$; $p < .01$; $\chi^2 = 11.08$; $p < .001$, respectively).

4. There were several indications that the inexperienced and expert groups differed in their structural patterns. Experts tended to make dots between dots longer, initial and dash-following terminal dashes shorter, and spaces between dots longer, relatively, than did the inexperienced.

Discussion

Bryan and Harter felt that heterotaxic variations resulted from operators attempting to develop distinctive styles by which they could be recognized (**1**, p. 34). More plausible explanations are: (*a*) that heterotaxic variations are a natural consequence of movement compounding, (*b*) that they are the consequence of attempts to imitate heard signals in which components of objectively equal duration appear to be of unequal duration, or (*c*) that they have been shaped by the reinforcing effects of successful imitation of signals heard in the past and of satisfactory reception by other operators.

It has been observed (**6, 12, 14**) that in skills like sending the operator achieves optimal efficiency by using "ballistic" movements, fast loose movements utilizing momentum, rather than slow forced movements. Such performance may be expected to require certain types of movement organizations. Stetson and Tuthill (**15**) found that the performance of musicians playing a single "dotted-eighth-sixteenth" phrase on a key did not correspond to the theoretical value of the notation. Since "the small variations show a precise movement type for each sitting," the investigators conceived of the "rhythmic unit-group as an organized group of movements whose conditions are muscular," wherein "changes of stress and of muscle group alter the type of unit-group" (**15**). The results of the present investigation are similar to those obtained by Stetson and Tuthill in revealing idiosyncratic discrepancies between the actual and the theoretical temporal relations. The fact that dot durations depended upon the kind of element preceded and followed argues that movement compounding results in heterotaxic variations.

Evidence of a relationship between perceptual distortions of signals and heterotaxic variations exists in studies of errors in receiving by Spragg (**13**), and by Seashore and Kurtz (**11**). Both of these studies emphasized a tendency for operators to shorten signals at the end. The general sending tendency, especially strong in the inexperienced operators, for the terminal dash to be elongated is consistent with the view that the terminal component must be longer than other components to sound equal to them. This view is supported by the literature on time-order errors in time estimation (**16**). If two consecutive equal intervals are as short as .90 sec., the second interval seems shorter than the first.

The fact that heterotaxic variations typify sending should be of considerable interest and importance for telegraphy. The most obvious relationship is that between different types or degrees of heterotaxic sending and receptability. Herbert and Keller (**7**) have shown that the use of hand-sent clear text in training may result in a definite improvement over the use of machine-sent cipher in learning to receive machine-sent cipher material. Although Herbert and Keller attributed this effect primarily to the use

of clear text, some benefit may have accrued from the exposure to heterotaxic variations in components.

References

1. Bryan, W. L., & Harter, N. Studies in the physiology and psychology of the telegraphic language. *Psychol. Rev.*, 1897, **4**, 27–53.
2. Crump, S. L. The estimation of variance components in analysis of variance. *Biometrics Bull.*, 1946, **2**, 7–11.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Flik, G. Ergebnisse einer experimentellen Untersuchung über die Wiedergabe der Fünkzeichen. *Z. angew. Psychol.*, 1938, **55**, 222–237.
5. Flik, G. Psychische Vorgänge beim Funken, im besondern die subjektiven Umformugen. *Z. Psychol.*, 1941, **150**, 320–383; **151**, 193–229.
6. Hartson, L. D. Analysis of skilled movements. *Person. J.*, 1932, **11**, 28–43.
7. Herbert, M. J., & Keller, F. S. *The use of hand-sent clear text as practice material in learning to receive Morse code.* (OSRD Report No. 5168, 1945: Publ. Bd., No. 12158) Washington, D. C.: U. S. Dep. Commerce, 1946.
8. Kurtz, A. K., Seashore, H., Stuntz, S., & Willits, J. M. *The standardization of code speeds.* (OSRD Report No. 3490, 1944: Publ. Bd., No. 12019) Washington, D. C.: U. S. Dep. Commerce, 1946.
9. Morsh, J. E., & Stannard, A. F. B. Studies in international Morse code. II. A simplified method of determining code speed. *Canad. J. Psychol.*, 1947, **1**, 67–70.
10. Morsh, J. E., & Stannard, A. F. B. Studies in international Morse code. VI. Speed calculations for international Morse code. *Canad. J. Psychol.*, 1948, **2**, 62–70.
11. Seashore, H., & Kurtz, A. K. *Analysis of errors in copying code.* (OSRD Report No. 4010, 1944: Publ. Bd., No. 12170) Washington, D. C.: U. S. Dep. Commerce, 1946.
12. Smith, M., Culpin, M., & Farmer, E. *A study of telegraphist's cramp.* Great Britain Med. Res. Council, Industrial Fatigue Research Board, Rep. No. 43, 1927. Pp. 40.
13. Spragg, S. D. S. The relative difficulty of Morse code alphabet characters learned by the whole method. *J. exp. Psychol.*, 1943, **33**, 108–114.
14. Stetson, R. H., & McDill, J. A. Mechanism of the different types of movement. *Psychol. Monogr.*, 1923, **32**, No. 3 (Whole No. 144).
15. Stetson, R. H., & Tuthill, T. E. Measurements of rhythmic unit-groups at different tempos. *Psychol. Monogr.*, 1923, **32**, No. 3 (Whole No. 144).
16. Weitz, J., & Fair, K. L. A survey of studies of rhythm. *USAF Hum. Resour. Res. Cent., Res. Bull.*, 1951, No. 51–4.

(Received September 29, 1954)